

162249

STIC-Biotech/ChemLib

From: Swope, Sheridan  
Sent: Thursday, August 11, 2005 12:01 PM  
To: STIC-Biotech/ChemLib  
Subject: 10/726,967

For 10/726,967, pls search:

SID 84 against the NT and AA data bases.

Thanks

Sheridan Swope, Ph.D.  
Patent Examiner, AU 1656  
Recombinant Enzymes  
571-272-0943 (voice)  
E02B71 Remsen Bld (Office)  
E03C70 Remsen Bld (Mailbox)

CRFE

Paula Shoppard

\*\*\*\*\*

STAFF USE ONLY

Searcher: \_\_\_\_\_  
Searcher Phone: 2-\_\_\_\_\_  
Date Searcher Picked up: \_\_\_\_\_  
Date Completed: \_\_\_\_\_  
Searcher Prep/Rev. Time: \_\_\_\_\_  
Online Time: \_\_\_\_\_

\*\*\*\*\*

Type of Search

NA#: \_\_\_\_\_ AA#: \_\_\_\_\_  
Interference: \_\_\_\_\_ SPDI: \_\_\_\_\_  
S/L: \_\_\_\_\_ Oligomer: \_\_\_\_\_  
Encode/Transl: \_\_\_\_\_  
Structure#: \_\_\_\_\_ Text: \_\_\_\_\_  
Inventor: \_\_\_\_\_ Litigation: \_\_\_\_\_

\*\*\*\*\*

Vendors and cost where applicable

STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
QUESTEL/ORBIT: \_\_\_\_\_  
LEXIS/NEXIS: \_\_\_\_\_  
SEQUENCE SYSTEM: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other(Specify): \_\_\_\_\_

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 21, 2005, 08:05:28 ; Search time 132 Seconds  
(without alignment)

1679.775 Million cell updates/sec

Title: US-10-726-967a-84

Perfect score: 2288

Sequence: 1 TQHGIRLPRLRSGLGAPLGL.....FVTLMDPCGNIPQTBEST 433

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Uniprot\_03:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2267	99.1	501	1 BAE1_HUMAN	P56817 homo sapien
2	2261	98.8	501	2 Q81YC8	Q81YC8 homo sapien
3	2218	96.9	501	1 BAE1_RAT	P56819 r beta-secr
4	2216	96.9	501	1 BAE1_MOUSE	P56818 mus musculu
5	2215	96.8	501	2 Q8C7R1	Q8C7R1 mus musculu
6	2211	96.6	501	2 Q8B0Y4	Q8B0Y4 mus musculu
7	2204	96.3	532	2 Q9UJ51	Q9UJ51 homo sapien
8	2023	88.4	467	2 Q8C4F4	Q8C4F4 mus musculu
9	1794	78.4	505	2 Q6NZT7	Q6NZT7 brachydanto
10	1711	74.8	373	2 Q69ZQ6	Q69ZQ6 mus musculu
11	1163	50.8	267	2 Q9CUT5	Q9CUT5 mus musculu
12	1136	49.7	518	1 BAE2_HUMAN	Q9H2V8 homo sapien
13	1121	49.0	439	2 Q9H2V8	Q9H2V8 homo sapien
14	1120	49.0	514	2 Q8C5B9	Q8C5B9 mus musculu
15	1120	49.0	514	2 Q9JUL8	Q9JUL8 mus musculu
16	1117	48.8	514	2 Q6IE75	Q6IE75 rattus norv
17	1115	48.7	514	2 Q8C7R3	Q8C7R3 mus musculu
18	1070	46.8	499	2 Q6B820	Q6B820 xenopus lae
19	1061	46.4	500	2 Q7T0Y2	Q7T0Y2 xenopus lae
20	1038	45.4	423	2 Q8N2D4	Q8N2D4 homo sapien
21	971.5	42.5	396	2 Q9NZL1	Q9NZL1 homo sapien
22	937	41.0	468	2 Q9NZL2	Q9NZL2 homo sapien
23	712.5	31.1	213	2 Q9PDD2	Q9PDD2 homo sapien
24	566.5	23.2	255	2 Q9R1P7	Q9R1P7 mus musculu
25	530	23.2	127	2 Q76KP0	Q76KP0 homo sapien
26	359.5	15.1	244	2 Q8WQY9	Q8WQY9 aphrocallis
27	345	15.1	76	2 Q8N698	Q8N698 homo sapien
28	336.5	14.7	392	2 Q6P7F8	Q6P7F8 xenopus lae
29	332.5	14.5	391	2 Q9VKP6	Q9VKP6 drosophila
30	332	14.5	354	2 Q9GTX7	Q9GTX7 boophilus m
31	327	14.3	324	1 PEP1_GADMO	P56272 gadus morhu

#### ALIGNMENTS

32	314.5	13.7	390	1	CATD_BOVIN	P80209 bos taurus
33	314	13.7	397	1	CTEB2_XENLA	Q805f2 xenopus lae
34	312.5	13.7	386	2	Q9BGU5	Q9BGU5 bos taurus
35	311	13.6	387	1	PEP1_RABIT	P28712 oryctolagus
36	309	13.5	384	2	Q9DEC2	Q9DEC2 xenopus lae
37	308	13.5	385	2	Q9DEC4	Q9DEC4 rana carseb
38	308	13.5	397	1	CTE1_XENLA	Q805f3 xenopus lae
39	307.5	13.4	401	2	Q6INM6	Q6INM6 xenopus lae
40	307	13.4	386	2	Q6HA02	Q6HA02 brachydanto
41	306	13.4	412	2	Q6GT7	Q6GT7 rhinolophus
42	305	13.3	386	2	Q9GMY7	Q9GMY7 sorex ungui
43	305	13.3	387	2	Q9GMY8	Q9GMY8 drosophila
44	304.5	13.3	372	2	Q9VTK3	Q9VTK3 drosophila
45	304.5	13.3	383	2	Q9DEC3	Q9DEC3 xenopus lae

#### RESULT 1

AC	BAE1_HUMAN	STANDARD;	PRT;	501 AA.
DT	30-MAY-2000	(Rel. 39, Created)		
DT	30-MAY-2000	(Rel. 39, Last sequence update)		
DT	25-OCT-2004	(Rel. 45, Last annotation update)		
DE	Beta-secretase 1 precursor (EC 3.4.23.46) (Beta-site APP cleaving enzyme 1) (Beta-site amyloid precursor protein cleaving enzyme 1) (Aspartyl protease 2) (Asp 2) (ASP2) (Membrane-associated aspartic protease 2) (Memapsin-2).			
DE	Name=BACE1; Synonyms=BACE;			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORM A).			
RC	TISSUE=Brain;			
RX	MEDLINE=20002972; PubMed=10531052; DOI=10.1126/science.286.5440.735;			
RA	Vassar R., Bennett B.D., Babu-Khan S., Kahn S., Mendiaz E.A.,			
RA	Denis P., Teplow D.B., Ross S., Amarante P., Loebf M., Luo Y.,			
RA	Fisher S., Fuller J., Edenson S., Lile J., Jarosinski M.A.,			
RA	Biere A.L., Curran E., Burgess T., Louis J.C., Collins F.,			
RA	Treanor J., Rogers G., Citron M.,			
RT	"Beta-secretase cleavage of Alzheimer's amyloid precursor protein by the transmembrane aspartic protease BACE.";			
RL	Science 286:735-741(1999).			
RN	[2]			
RP	SEQUENCE FROM N.A. (ISOFORM A), SEQUENCE OF 46-68, AND CHARACTERIZATION.			
RC	TISSUE=Brain;			
RX	MEDLINE=20057171; PubMed=10591214; DOI=10.1038/990114;			
RA	Sinha S., Anderson J.P., Barbour R., Basl G.S., Caccavello R.,			
RA	Davis D., Doan M., Dovey H.F., Frigon N., Hong J., Jacobson-Croak K.,			
RA	Jewett N., Keim P., Knops J., Lieberburg I., Power M., Tan H.,			
RA	Tateo D., Zhao J., Schenk D., Seubert P., Suematsu S.M., Wang S.,			
RA	Walker D., Zhao J., McConlogue L., Varghese J.,			
RT	"Purification and cloning of amyloid precursor protein beta-secretase from human brain.";			
RL	Nature 402:537-540(1999).			
RN	[3]			
RP	SEQUENCE FROM N.A. (ISOFORM A)			
RX	MEDLINE=20057170; PubMed=10591213; DOI=10.1038/990107;			
RA	Yan R., Bienkowski M.J., Shuck M.E., Miao H., Tory M.C., Pauley A.M.,			
RA	Braslier J.R., Strahman N.C., Mathews W.R., Buhl A.E., Carter D.B.,			
RA	Tomasek A.G., Parodi L.A., Heinkeon R.L., Gurney M.E.,			
RT	"Membrane-anchored aspartyl protease with Alzheimer's disease beta-secretase activity.";			
RL	Nature 402:533-537(1999).			
RN	[4]			
RP	SEQUENCE FROM N.A. (ISOFORM A).			
RX	MEDLINE=20120043; PubMed=10656250; DOI=10.1006/mcne.1999.0811;			
RA	Hussein I., Powell D.J., Howlett D.R., Tew D.G., Meek T.D.,			

RA Chapman C., Gloger I.S., Murphy K.E., Southan C.D., Ryan D.M.,  
 RA Smith T.S., Simons D.L., Walsh F.S., Dingwall C., Christie G.,  
 RT "Identification of a novel aspartic proteolase (Asp 2) as beta-  
 RT secretase.";  
 RL Mol. Cell. Neurosci. 14:419-427(1999).  
 RN (5)  
 RP SEQUENCE FROM N.A. (ISOFORM B).  
 RC TISSUE=Brain, and Pancreas;  
 RA Michel B., De Pietri Tonelli D., Zaccchetti D., Keller P.;  
 RT "New beta-site APP cleaving enzyme isoform (BACE-1B) obtained from  
 RT human brain and pancreas.";  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 RN (6)  
 RP SEQUENCE FROM N.A. (ISOFORM C).  
 RC TISSUE=Pancreas;  
 RA Zaccchetti D., De Pietri Tonelli D., Schnurbus R.;  
 RT "New beta-site APP cleaving enzyme isoform (BACE-1C) obtained from  
 RT human pancreas.";  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 RN (7)  
 RP SEQUENCE FROM N.A. (ISOFORMS B; C AND D).  
 RC TISSUE=Brain;  
 RX MEDLINE=21408467; PubMed=11516562; DOI=10.1016/S0304-3940(01)01912-7;  
 RA Tanahashi H., Tabira T.;  
 RT "Three novel alternatively spliced isoforms of the human beta-site  
 RT amyloid precursor protein cleaving enzyme (BACE) and their effect on  
 RL Neurosci. Lett. 307:9-12(2001).  
 RN (8)  
 RP SEQUENCE OF 14-501 FROM N.A. (ISOFORM A), AND CHARACTERIZATION.  
 RX MEDLINE=2014060; PubMed=10677483; DOI=10.1073/pnas.97.4.1456;  
 RA Lin X., Koelsch G., Wu S., Downs D., Dashi A., Tang J.;  
 RT "Human aspartic protease memapsin 2 cleaves the beta-secretase site of  
 RT beta-amyloid precursor protein.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 97:1456-1460(2000).  
 RN (9)  
 RP DISULFIDE BONDS.  
 RX MEDLINE=21950860; PubMed=11953458;  
 RA Fischer F., Molinari M., Bodendorf U., Paganetti P.;  
 RT "The disulphide bonds in the catalytic domain of BACE are critical but  
 RT not essential for amyloid precursor protein processing activity.";  
 RL J. Neurochem. 80:1079-1088(2002).  
 RN (10)  
 RP FUNCTION: Responsible for the proteolytic processing of the  
 CC amyloid precursor protein (APP). Cleaves at the amino terminus of  
 CC the A-beta peptide sequence, between residues 671 and 672 of APP,  
 CC leads to the generation and extracellular release of beta-cleaved  
 CC soluble APP, and a corresponding cell-associated carboxy-terminal  
 CC fragment which is later released by gamma-secretase.  
 CC -1- CATALYTIC ACTIVITY: Broad endopeptidase specificity. Cleaves Glu-  
 CC Val-Aen-Leu-|-Asp-Ala-Glu-Phe in the Swedish variant of  
 CC Alzheimer's amyloid precursor protein.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Name=A; Synonyms=BACE-1A, BAC-501;  
 CC Name=B; Synonyms=BACE-1B, BACE-1-476;  
 CC Name=C; Synonyms=BACE-1C, BACE-1-457;  
 CC Name=D; Synonyms=BACE-1D, BACE-1-432;  
 CC Name=E; Synonyms=BACE-1E, BACE-1-433;  
 CC Name=F; Synonyms=BACE-1F, BACE-1-434;  
 CC Name=G; Synonyms=BACE-1G, BACE-1-435;  
 CC Name=H; Synonyms=BACE-1H, BACE-1-436;  
 CC Name=I; Synonyms=BACE-1I, BACE-1-437;  
 CC Name=J; Synonyms=BACE-1J, BACE-1-438;  
 CC Name=K; Synonyms=BACE-1K, BACE-1-439;  
 CC Name=L; Synonyms=BACE-1L, BACE-1-440;  
 CC Name=M; Synonyms=BACE-1M, BACE-1-441;  
 CC Name=N; Synonyms=BACE-1N, BACE-1-442;  
 CC Name=O; Synonyms=BACE-1O, BACE-1-443;  
 CC Name=P; Synonyms=BACE-1P, BACE-1-444;  
 CC Name=Q; Synonyms=BACE-1Q, BACE-1-445;  
 CC Name=R; Synonyms=BACE-1R, BACE-1-446;  
 CC Name=S; Synonyms=BACE-1S, BACE-1-447;  
 CC Name=T; Synonyms=BACE-1T, BACE-1-448;  
 CC Name=U; Synonyms=BACE-1U, BACE-1-449;  
 CC Name=V; Synonyms=BACE-1V, BACE-1-450;  
 CC Name=W; Synonyms=BACE-1W, BACE-1-451;  
 CC Name=X; Synonyms=BACE-1X, BACE-1-452;  
 CC Name=Y; Synonyms=BACE-1Y, BACE-1-453;  
 CC Name=Z; Synonyms=BACE-1Z, BACE-1-454;  
 CC Name=AA; Synonyms=BACE-1AA, BACE-1-455;  
 CC Name=AB; Synonyms=BACE-1AB, BACE-1-456;  
 CC Name=AC; Synonyms=BACE-1AC, BACE-1-457;  
 CC Name=AD; Synonyms=BACE-1AD, BACE-1-458;  
 CC Name=AE; Synonyms=BACE-1AE, BACE-1-459;  
 CC Name=AF; Synonyms=BACE-1AF, BACE-1-460;  
 CC Name=AG; Synonyms=BACE-1AG, BACE-1-461;  
 CC Name=AH; Synonyms=BACE-1AH, BACE-1-462;  
 CC Name=AI; Synonyms=BACE-1AI, BACE-1-463;  
 CC Name=AJ; Synonyms=BACE-1AJ, BACE-1-464;  
 CC Name=AK; Synonyms=BACE-1AK, BACE-1-465;  
 CC Name=AL; Synonyms=BACE-1AL, BACE-1-466;  
 CC Name=AM; Synonyms=BACE-1AM, BACE-1-467;  
 CC Name=AN; Synonyms=BACE-1AN, BACE-1-468;  
 CC Name=AO; Synonyms=BACE-1AO, BACE-1-469;  
 CC Name=AP; Synonyms=BACE-1AP, BACE-1-470;  
 CC Name=AQ; Synonyms=BACE-1AQ, BACE-1-471;  
 CC Name=AR; Synonyms=BACE-1AR, BACE-1-472;  
 CC Name=AS; Synonyms=BACE-1AS, BACE-1-473;  
 CC Name=AT; Synonyms=BACE-1AT, BACE-1-474;  
 CC Name=AU; Synonyms=BACE-1AU, BACE-1-475;  
 CC Name=AV; Synonyms=BACE-1AV, BACE-1-476;  
 CC Name=AW; Synonyms=BACE-1AW, BACE-1-477;  
 CC Name=AX; Synonyms=BACE-1AX, BACE-1-478;  
 CC Name=AY; Synonyms=BACE-1AY, BACE-1-479;  
 CC Name=AZ; Synonyms=BACE-1AZ, BACE-1-480;  
 CC Name=BA; Synonyms=BACE-1BA, BACE-1-481;  
 CC Name=BB; Synonyms=BACE-1BB, BACE-1-482;  
 CC Name=BC; Synonyms=BACE-1BC, BACE-1-483;  
 CC Name=BD; Synonyms=BACE-1BD, BACE-1-484;  
 CC Name=BE; Synonyms=BACE-1BE, BACE-1-485;  
 CC Name=BF; Synonyms=BACE-1BF, BACE-1-486;  
 CC Name=BG; Synonyms=BACE-1BG, BACE-1-487;  
 CC Name=BG; Synonyms=BACE-1BG, BACE-1-488;  
 CC Name=BI; Synonyms=BACE-1BI, BACE-1-489;  
 CC Name=BJ; Synonyms=BACE-1BJ, BACE-1-490;  
 CC Name=BK; Synonyms=BACE-1BK, BACE-1-491;  
 CC Name=BL; Synonyms=BACE-1BL, BACE-1-492;  
 CC Name=BM; Synonyms=BACE-1BM, BACE-1-493;  
 CC Name=BN; Synonyms=BACE-1BN, BACE-1-494;  
 CC Name=BO; Synonyms=BACE-1BO, BACE-1-495;  
 CC Name=BP; Synonyms=BACE-1BP, BACE-1-496;  
 CC Name=BQ; Synonyms=BACE-1BQ, BACE-1-497;  
 CC Name=BR; Synonyms=BACE-1BR, BACE-1-498;  
 CC Name=BS; Synonyms=BACE-1BS, BACE-1-499;  
 CC Name=BT; Synonyms=BACE-1BT, BACE-1-500;  
 CC Name=BU; Synonyms=BACE-1BU, BACE-1-501;  
 CC Name=BV; Synonyms=BACE-1BV, BACE-1-502;  
 CC Name=BW; Synonyms=BACE-1BW, BACE-1-503;  
 CC Name=BX; Synonyms=BACE-1BX, BACE-1-504;  
 CC Name=BY; Synonyms=BACE-1BY, BACE-1-505;  
 CC Name=BZ; Synonyms=BACE-1BZ, BACE-1-506;  
 CC Name=CA; Synonyms=BACE-1CA, BACE-1-507;  
 CC Name=CB; Synonyms=BACE-1CB, BACE-1-508;  
 CC Name=CC; Synonyms=BACE-1CC, BACE-1-509;  
 CC Name=CD; Synonyms=BACE-1CD, BACE-1-510;  
 CC Name=CE; Synonyms=BACE-1CE, BACE-1-511;  
 CC Name=CF; Synonyms=BACE-1CF, BACE-1-512;  
 CC Name=CG; Synonyms=BACE-1CG, BACE-1-513;  
 CC Name=CH; Synonyms=BACE-1CH, BACE-1-514;  
 CC Name=CI; Synonyms=BACE-1CI, BACE-1-515;  
 CC Name=CJ; Synonyms=BACE-1CJ, BACE-1-516;  
 CC Name=CK; Synonyms=BACE-1CK, BACE-1-517;  
 CC Name=CL; Synonyms=BACE-1CL, BACE-1-518;  
 CC Name=CM; Synonyms=BACE-1CM, BACE-1-519;  
 CC Name=CN; Synonyms=BACE-1CN, BACE-1-520;  
 CC Name=CO; Synonyms=BACE-1CO, BACE-1-521;  
 CC Name=CP; Synonyms=BACE-1CP, BACE-1-522;  
 CC Name=CQ; Synonyms=BACE-1CQ, BACE-1-523;  
 CC Name=CR; Synonyms=BACE-1CR, BACE-1-524;  
 CC Name=CS; Synonyms=BACE-1CS, BACE-1-525;  
 CC Name=CT; Synonyms=BACE-1CT, BACE-1-526;  
 CC Name=CU; Synonyms=BACE-1CU, BACE-1-527;  
 CC Name=CV; Synonyms=BACE-1CV, BACE-1-528;  
 CC Name=CW; Synonyms=BACE-1CW, BACE-1-529;  
 CC Name=CX; Synonyms=BACE-1CX, BACE-1-530;  
 CC Name=CY; Synonyms=BACE-1CY, BACE-1-531;  
 CC Name=CA; Synonyms=BACE-1CA, BACE-1-532;  
 CC Name=CB; Synonyms=BACE-1CB, BACE-1-533;  
 CC Name=CC; Synonyms=BACE-1CC, BACE-1-534;  
 CC Name=CD; Synonyms=BACE-1CD, BACE-1-535;  
 CC Name=CE; Synonyms=BACE-1CE, BACE-1-536;  
 CC Name=CF; Synonyms=BACE-1CF, BACE-1-537;  
 CC Name=CG; Synonyms=BACE-1CG, BACE-1-538;  
 CC Name=CH; Synonyms=BACE-1CH, BACE-1-539;  
 CC Name=CI; Synonyms=BACE-1CI, BACE-1-540;  
 CC Name=CJ; Synonyms=BACE-1CJ, BACE-1-541;  
 CC Name=CK; Synonyms=BACE-1CK, BACE-1-542;  
 CC Name=CL; Synonyms=BACE-1CL, BACE-1-543;  
 CC Name=CM; Synonyms=BACE-1CM, BACE-1-544;  
 CC Name=CN; Synonyms=BACE-1CN, BACE-1-545;  
 CC Name=CO; Synonyms=BACE-1CO, BACE-1-546;  
 CC Name=CP; Synonyms=BACE-1CP, BACE-1-547;  
 CC Name=CQ; Synonyms=BACE-1CQ, BACE-1-548;  
 CC Name=CR; Synonyms=BACE-1CR, BACE-1-549;  
 CC Name=CS; Synonyms=BACE-1CS, BACE-1-550;  
 CC Name=CT; Synonyms=BACE-1CT, BACE-1-551;  
 CC Name=CU; Synonyms=BACE-1CU, BACE-1-552;  
 CC Name=CV; Synonyms=BACE-1CV, BACE-1-553;  
 CC Name=CW; Synonyms=BACE-1CW, BACE-1-554;  
 CC Name=CX; Synonyms=BACE-1CX, BACE-1-555;  
 CC Name=CY; Synonyms=BACE-1CY, BACE-1-556;  
 CC Name=CA; Synonyms=BACE-1CA, BACE-1-557;  
 CC Name=CB; Synonyms=BACE-1CB, BACE-1-558;  
 CC Name=CC; Synonyms=BACE-1CC, BACE-1-559;  
 CC Name=CD; Synonyms=BACE-1CD, BACE-1-560;  
 CC Name=CE; Synonyms=BACE-1CE, BACE-1-561;  
 CC Name=CF; Synonyms=BACE-1CF, BACE-1-562;  
 CC Name=CG; Synonyms=BACE-1CG, BACE-1-563;  
 CC Name=CH; Synonyms=BACE-1CH, BACE-1-564;  
 CC Name=CI; Synonyms=BACE-1CI, BACE-1-565;  
 CC Name=CJ; Synonyms=BACE-1CJ, BACE-1-566;  
 CC Name=CK; Synonyms=BACE-1CK, BACE-1-567;  
 CC Name=CL; Synonyms=BACE-1CL, BACE-1-568;  
 CC Name=CM; Synonyms=BACE-1CM, BACE-1-569;  
 CC Name=CN; Synonyms=BACE-1CN, BACE-1-570;  
 CC Name=CO; Synonyms=BACE-1CO, BACE-1-571;  
 CC Name=CP; Synonyms=BACE-1CP, BACE-1-572;  
 CC Name=CQ; Synonyms=BACE-1CQ, BACE-1-573;  
 CC Name=CR; Synonyms=BACE-1CR, BACE-1-574;  
 CC Name=CS; Synonyms=BACE-1CS, BACE-1-575;  
 CC Name=CT; Synonyms=BACE-1CT, BACE-1-576;  
 CC Name=CU; Synonyms=BACE-1CU, BACE-1-577;  
 CC Name=CV; Synonyms=BACE-1CV, BACE-1-578;  
 CC Name=CW; Synonyms=BACE-1CW, BACE-1-579;  
 CC Name=CX; Synonyms=BACE-1CX, BACE-1-580;  
 CC Name=CY; Synonyms=BACE-1CY, BACE-1-581;  
 CC Name=CA; Synonyms=BACE-1CA, BACE-1-582;  
 CC Name=CB; Synonyms=BACE-1CB, BACE-1-583;  
 CC Name=CC; Synonyms=BACE-1CC, BACE-1-584;  
 CC Name=CD; Synonyms=BACE-1CD, BACE-1-585;  
 CC Name=CE; Synonyms=BACE-1CE, BACE-1-586;  
 CC Name=CF; Synonyms=BACE-1CF, BACE-1-587;  
 CC Name=CG; Synonyms=BACE-1CG, BACE-1-588;  
 CC Name=CH; Synonyms=BACE-1CH, BACE-1-589;  
 CC Name=CI; Synonyms=BACE-1CI, BACE-1-590;  
 CC Name=CJ; Synonyms=BACE-1CJ, BACE-1-591;  
 CC Name=CK; Synonyms=BACE-1CK, BACE-1-592;  
 CC Name=CL; Synonyms=BACE-1CL, BACE-1-593;  
 CC Name=CM; Synonyms=BACE-1CM, BACE-1-594;  
 CC Name=CN; Synonyms=BACE-1CN, BACE-1-595;  
 CC Name=CO; Synonyms=BACE-1CO, BACE-1-596;  
 CC Name=CP; Synonyms=BACE-1CP, BACE-1-597;  
 CC Name=CQ; Synonyms=BACE-1CQ, BACE-1-598;  
 CC Name=CR; Synonyms=BACE-1CR, BACE-1-599;  
 CC Name=CS; Synonyms=BACE-1CS, BACE-1-600;  
 CC Name=CT; Synonyms=BACE-1CT, BACE-1-601;  
 CC Name=CU; Synonyms=BACE-1CU, BACE-1-602;  
 CC Name=CV; Synonyms=BACE-1CV, BACE-1-603;  
 CC Name=CW; Synonyms=BACE-1CW, BACE-1-604;  
 CC Name=CX; Synonyms=BACE-1CX, BACE-1-605;  
 CC Name=CY; Synonyms=BACE-1CY, BACE-1-606;  
 CC Name=CA; Synonyms=BACE-1CA, BACE-1-607;  
 CC Name=CB; Synonyms=BACE-1CB, BACE-1-608;  
 CC Name=CC; Synonyms=BACE-1CC, BACE-1-609;  
 CC Name=CD; Synonyms=BACE-1CD, BACE-1-610;  
 CC Name=CE; Synonyms=BACE-1CE, BACE-1-611;  
 CC Name=CF; Synonyms=BACE-1CF, BACE-1-612;  
 CC Name=CG; Synonyms=BACE-1CG, BACE-1-613;  
 CC Name=CH; Synonyms=BACE-1CH, BACE-1-614;  
 CC Name=CI; Synonyms=BACE-1CI, BACE-1-615;  
 CC Name=CJ; Synonyms=BACE-1CJ, BACE-1-616;  
 CC Name=CK; Synonyms=BACE-1CK, BACE-1-617;  
 CC Name=CL; Synonyms=BACE-1CL, BACE-1-618;  
 CC Name=CM; Synonyms=BACE-1CM, BACE-1-619;  
 CC Name=CN; Synonyms=BACE-1CN, BACE-1-620;  
 CC Name=CO; Synonyms=BACE-1CO, BACE-1-621;  
 CC Name=CP; Synonyms=BACE-1CP, BACE-1-622;  
 CC Name=CQ; Synonyms=BACE-1CQ, BACE-1-623;  
 CC Name=CR; Synonyms=BACE-1CR, BACE-1-624;  
 CC Name=CS; Synonyms=BACE-1CS, BACE-1-625;  
 CC Name=CT; Synonyms=BACE-1CT, BACE-1-626;  
 CC Name=CU; Synonyms=BACE-1CU, BACE-1-627;  
 CC Name=CV; Synonyms=BACE-1CV, BACE-1-628;  
 CC Name=CW; Synonyms=BACE-1CW, BACE-1-629;  
 CC Name=CX; Synonyms=BACE-1CX, BACE-1-630;  
 CC Name=CY; Synonyms=BACE-1CY, BACE-1-631;  
 CC Name=CA; Synonyms=BACE-1CA, BACE-1-632;  
 CC Name=CB; Synonyms=BACE-1CB, BACE-1-633;  
 CC Name=CC; Synonyms=BACE-1CC, BACE-1-634;  
 CC Name=CD; Synonyms=BACE-1CD, BACE-1-635;  
 CC Name=CE; Synonyms=BACE-1CE, BACE-1-636;  
 CC Name=CF; Synonyms=BACE-1CF, BACE-1-637;  
 CC Name=CG; Synonyms=BACE-1CG, BACE-1-638;  
 CC Name=CH; Synonyms=BACE-1CH, BACE-1-639;  
 CC Name=CI; Synonyms=BACE-1CI, BACE-1-640;  
 CC Name=CJ; Synonyms=BACE-1CJ, BACE-1-641;  
 CC Name=CK; Synonyms=BACE-1CK, BACE-1-642;  
 CC Name=CL; Synonyms=BACE-1CL, BACE-1-643;  
 CC Name=CM; Synonyms=BACE-1CM, BACE-1-644;  
 CC Name=CN; Synonyms=BACE-1CN, BACE-1-645;  
 CC Name=CO; Synonyms=BACE-1CO, BACE-1-646;  
 CC Name=CP; Synonyms=BACE-1CP, BACE-1-647;  
 CC Name=CQ; Synonyms=BACE-1CQ, BACE-1-648;  
 CC Name=CR; Synonyms=BACE-1CR, BACE-1-649;  
 CC Name=CS; Synonyms=BACE-1CS, BACE-1-650;  
 CC Name=CT; Synonyms=BACE-1CT, BACE-1-651;  
 CC Name=CU; Synonyms=BACE-1CU, BACE-1-652;  
 CC Name=CV; Synonyms=BACE-1CV, BACE-1-653;  
 CC Name=CW; Synonyms=BACE-1CW, BACE-1-654;  
 CC Name=CX; Synonyms=BACE-1CX, BACE-1-655;  
 CC Name=CY; Synonyms=BACE-1CY, BACE-1-656;  
 CC Name=CA; Synonyms=BACE-1CA, BACE-1-657;  
 CC Name=CB; Synonyms=BACE-1CB, BACE-1-658;  
 CC Name=CC; Synonyms=BACE-1CC, BACE-1-659;  
 CC Name=CD; Synonyms=BACE-1CD, BACE-1-660;  
 CC Name=CE; Synonyms=BACE-1CE, BACE-1-661;  
 CC Name=CF; Synonyms=BACE-1CF, BACE-1-662;  
 CC Name=CG; Synonyms=BACE-1CG, BACE-1-663;  
 CC Name=CH; Synonyms=BACE-1CH, BACE-1-664;  
 CC Name=CI; Synonyms=BACE-1CI, BACE-1-665;  
 CC Name=CJ; Synonyms=BACE-1CJ, BACE-1-666;  
 CC Name=CK; Synonyms=BACE-1CK, BACE-1-667;  
 CC Name=CL; Synonyms=BACE-1CL, BACE-1-668;  
 CC Name=CM; Synonyms=BACE-1CM, BACE-1-669;  
 CC Name=CN; Synonyms=BACE-1CN, BACE-1-670;  
 CC Name=CO; Synonyms=BACE-1CO, BACE-1-671;  
 CC Name=CP; Synonyms=BACE-1CP, BACE-1-672;  
 CC Name=CQ; Synonyms=BACE-1CQ, BACE-1-673;  
 CC Name=CR; Synonyms=BACE-1CR, BACE-1-674;  
 CC Name=CS; Synonyms=BACE-1CS, BACE-1-675;  
 CC Name=CT; Synonyms=BACE-1CT, BACE-1-676;  
 CC Name=CU; Synonyms=BACE-1CU, BACE-1-677;  
 CC Name=CV; Synonyms=BACE-1CV, BACE-1-678;  
 CC Name=CW; Synonyms=BACE-1CW, BACE-1-679;  
 CC Name=CX; Synonyms=BACE-1CX, BACE-1-680;  
 CC Name=CY; Synonyms=BACE-1CY, BACE-1-681;  
 CC Name=CA; Synonyms=BACE-1CA, BACE-1-682;  
 CC Name=CB; Synonyms=BACE-1CB, BACE-1-683;  
 CC Name=CC; Synonyms=BACE-1CC, BACE-1-684;  
 CC Name=CD; Synonyms=BACE-1CD, BACE-1-685;  
 CC Name=CE; Synonyms=BACE-1CE, BACE-1-686;  
 CC Name=CF; Synonyms=BACE-1CF, BACE-1-687;  
 CC Name=CG; Synonyms=BACE-1CG, BACE-1-688;  
 CC Name=CH; Synonyms=BACE-1CH, BACE-1-689;  
 CC Name=CI; Synonyms=BACE-1CI, BACE-1-690;  
 CC Name=CJ; Synonyms=BACE-1CJ, BACE-1-691;  
 CC Name=CK; Synonyms=BACE-1CK, BACE-1-692;  
 CC Name=CL; Synonyms=BACE-1CL, BACE-1-693;  
 CC Name=CM; Synonyms=BACE-1CM, BACE-1-694;  
 CC Name=CN; Synonyms=BACE-1CN, BACE-1-695;  
 CC Name=CO; Synonyms=BACE-1CO, BACE-1-696;  
 CC Name=CP; Synonyms=BACE-1CP, BACE-1-697;  
 CC Name=CQ; Synonyms=BACE-1CQ, BACE-1-698;  
 CC Name=CR; Synonyms=BACE-1CR, BACE-1-699;  
 CC Name=CS; Synonyms=BACE-1CS, BACE-1-700;  
 CC Name=CT; Synonyms=BACE-1CT, BACE-1-701;  
 CC Name=CU; Synonyms=BACE-1CU, BACE-1-702;  
 CC Name=CV; Synonyms=BACE-1CV, BACE-1-703;  
 CC Name=CW; Synonyms=BACE-1CW, BACE-1-704;  
 CC Name=CX; Synonyms=BACE-1CX, BACE-1-705;  
 CC Name=CY; Synonyms=BACE-1CY, BACE-1-706;  
 CC Name=CA; Synonyms=BACE-1CA, BACE-1-707;  
 CC Name=CB; Synonyms=BACE-1CB, BACE-1-708;  
 CC Name=CC; Synonyms=BACE-1CC, BACE-1-709;  
 CC Name=CD; Synonyms=BACE-1CD, BACE-1-710;  
 CC Name=CE; Synonyms=BACE-1CE, BACE-1-711;  
 CC Name=CF; Synonyms=BACE-1CF, BACE-1-712;  
 CC Name=CG; Synonyms=BACE-1CG, BACE-1-713;  
 CC Name=CH; Synonyms=BACE-1CH, BACE-1-714;  
 CC Name=CI; Synonyms=BACE-1CI, BACE-1-715;  
 CC Name=CJ; Synonyms=BACE-1CJ, BACE-1-716;  
 CC Name=CK; Synonyms=BACE-1CK, BACE-1-717;  
 CC Name=CL; Synonyms=BACE-1CL, BACE-1-718;  
 CC Name=CM; Synonyms=BACE-1CM, BACE-1-719;  
 CC Name=CN; Synonyms=BACE-1CN, BACE-1-720;  
 CC Name=CO; Synonyms=BACE-1CO, BACE-1-721;  
 CC Name=CP; Synonyms=BACE-1CP, BACE-1-722;  
 CC Name=CQ; Synonyms=BACE-1CQ, BACE-1-723;  
 CC Name=CR; Synonyms=BACE-1CR, BACE-1-724;  
 CC Name=CS; Synonyms=BACE-1CS, BACE-1-725;  
 CC Name=CT; Synonyms=BACE-1CT, BACE-1-726;  
 CC Name=CU; Synonyms=BACE-1CU, BACE-1-727;  
 CC Name=CV; Synonyms=BACE-1CV, BACE-1-728;  
 CC Name=CW; Synonyms=BACE-1CW, BACE-1-729;  
 CC Name=CX; Synonyms=BACE-1CX, BACE-1-730;  
 CC Name=CY; Synonyms=BACE-1CY, BACE-1-731;  
 CC Name=CA; Synonyms=BACE-1CA, BACE-1-732;  
 CC Name=CB; Synonyms=BACE-1CB, BACE-1-733;  
 CC Name=CC; Synonyms=BACE-1CC, BACE-1-734;  
 CC Name=CD; Synonyms=BACE-1CD, BACE-1-735;  
 CC Name=CE; Synonyms=BACE-1CE, BACE-1-736;  
 CC Name=CF; Synonyms=BACE-1CF, BACE-1-737;  
 CC Name=CG; Synonyms=BACE-1CG, BACE-1-738;  
 CC Name=CH; Synonyms=BACE-1CH, BACE-1-739;  
 CC Name=CI; Synonyms=BACE-1CI, BACE-1-740;  
 CC Name=CJ; Synonyms=BACE-1CJ, BACE-1-741;  
 CC Name=CK; Synonyms=BACE-1CK, BACE-1-742;  
 CC Name=CL; Synonyms=BACE-1CL, BACE-1-743;  
 CC Name=CM; Synonyms=BACE-1CM, BACE-1-744;  
 CC Name=CN; Synonyms=BACE-1CN, BACE-1-745;  
 CC Name=CO; Synonyms=BACE-1CO, BACE-1-746;  
 CC Name=CP; Synonyms=BACE-1CP, BACE-1-747;  
 CC Name=CQ; Synonyms=BACE-1CQ, BACE-1-748;  
 CC Name=CR; Synonyms=BACE-1CR, BACE-1-749;  
 CC Name=CS; Synonyms=BACE-1CS, BACE-1-750;  
 CC Name=CT; Synonyms=BACE-1CT, BACE-1-751;  
 CC Name=CU; Synonyms=BACE-1CU, BACE-1-752;  
 CC Name=CV; Synonyms=BACE-1CV, BACE-1-753;  
 CC Name=CW; Synonyms=BACE-1CW, BACE-1-754;  
 CC Name=CX; Synonyms=BACE-1CX, BACE-1-755;  
 CC Name=CY; Synonyms=BACE-1CY, BACE-1-756;  
 CC Name=CA; Synonyms=BACE-1CA, BACE-1-757;  
 CC Name=CB; Synonyms=BACE-1CB, BACE-1-758;  
 CC Name=CC; Synonyms=BACE-1CC, BACE-1-759;  
 CC Name=CD; Synonyms=BACE-1CD, BACE-1-760;  
 CC Name=CE; Synonyms=BACE-1CE, BACE-1-761;  
 CC Name=CF; Synonyms=BACE-1CF, BACE-1-762;  
 CC Name=CG; Synonyms=BACE-1CG, BACE-1-763;  
 CC Name=CH; Synonyms=BACE-1CH, BACE-1-764;  
 CC Name=CI; Synonyms=BACE-1CI, BACE-1-765;  
 CC Name=CJ; Synonyms=BACE-1CJ, BACE-1-766;  
 CC Name=CK; Synonyms=BACE-1CK, BACE-1-767;  
 CC Name=CL; Synonyms=BACE-1CL, BACE-1-768;  
 CC Name=CM; Synonyms=BACE-1CM, BACE-1-769;  
 CC Name=CN; Synonyms=BACE-1CN, BACE-1-770;  
 CC Name=CO; Synonyms=BACE-1CO, BACE-1-771;  
 CC Name=CP; Synonyms=BACE-1CP, BACE-1-772;  
 CC Name=CQ; Synonyms=BACE-1CQ, BACE-1-773;  
 CC Name=CR; Synonyms=BACE-1CR, BACE-1-774;  
 CC Name=CS; Synonyms=BACE-1CS, BACE-1-775;  
 CC Name=CT; Synonyms=BACE-1CT, BACE-1-776;  
 CC Name=CU; Synonyms=BACE-1CU, BACE-1-777;  
 CC Name=CV; Synonyms=BACE-1CV, BACE-1-778;  
 CC Name=CW; Synonyms=BACE-1CW, BACE-1-779;  
 CC Name=CX; Synonyms=BACE-1CX, BACE-1-780;  
 CC Name=CY; Synonyms=BACE-1CY, BACE-1-781;  
 CC Name=CA; Synonyms=BACE-1CA, BACE-1-782;  
 CC Name=CB; Synonyms=BACE-1CB, BACE-1-783;  
 CC Name=CC; Synonyms=BACE-1CC, BACE-1-784;  
 CC Name=CD; Synonyms=BACE-1CD, BACE-1-785;  
 CC Name=CE; Synonyms=BACE-1CE, BACE-1-786;  
 CC Name=CF; Synonyms=BACE-1CF, BACE-1-787;  
 CC Name=CG; Synonyms=BACE-1CG, BACE-1-788;  
 CC Name=CH; Synonyms=BACE-1CH, BACE-1-789;  
 CC Name=CI; Synonyms=BACE-1CI, BACE-1-790;  
 CC Name=CJ; Synonyms=BACE-1CJ, BACE-1-791;  
 CC Name=CK; Synonyms=BACE-1CK, BACE-1-792;  
 CC Name=CL; Synonyms=BACE-1CL, BACE-1-793;  
 CC Name=CM; Synonyms=BACE-1CM, BACE-1-794;  
 CC Name=CN; Synonyms=BACE-1CN, BACE-1-795;  
 CC Name=CO; Synonyms=BACE-1CO, BACE-1-796;  
 CC Name=CP; Synonyms=BACE-1CP, BACE-1-797;  
 CC Name=CQ; Synonyms=BACE-1CQ, BACE-1-798;  
 CC Name=CR; Synonyms=BACE-1CR, BACE-1-799;  
 CC Name=CS; Synonyms=BACE-1CS, BACE-1-800;  
 CC Name=CT; Synonyms=BACE-1CT, BACE-1-801;  
 CC Name=CU; Synonyms=BACE-1CU, BACE-1-802;  
 CC Name=CV; Synonyms=BACE-1CV, BACE-1-803;  
 CC Name=CW; Synonyms=BACE-1CW, BACE-1-804;  
 CC Name=CX; Synonyms=BACE-1CX, BACE-1-805;  
 CC Name=CY; Synonyms=BACE-1CY, BACE-1-806;  
 CC Name=CA; Synonyms=BACE-1CA, BACE-1-807;  
 CC Name=CB; Synonyms=BACE-1CB, BACE-1-808;  
 CC Name=CC; Synonyms=BACE-1CC, BACE-1-809;  
 CC Name=CD; Synonyms=BACE-1CD, BACE-1-810;  
 CC Name=CE; Synonyms=BACE-1CE, BACE-1-811;  
 CC Name=CF; Synonyms=BACE-1CF, BACE-1-812;  
 CC Name=CG; Synonyms=BACE-1CG, BACE-1-813;  
 CC Name=CH; Synonyms=BACE-1CH, BACE-1-814;  
 CC Name=CI; Synonyms=BACE-1CI, BACE-1-815;  
 CC Name=CJ; Synonyms=BACE-1CJ, BACE-1-816;  
 CC Name=CK; Synonyms=BACE-1CK, BACE-1-817;  
 CC Name=CL; Synonyms=BACE-1CL, BACE-1-818;  
 CC Name=CM; Synonyms=BACE-1CM, BACE-1-819;  
 CC Name=CN; Synonyms=BACE-1CN, BACE-1-820;  
 CC Name=CO; Synonyms=BACE-1CO, BACE-1-821;  
 CC Name=CP; Synonyms=BACE-1CP, BACE-1-822;  
 CC Name=CQ; Synonyms=BACE-1CQ, BACE-1-823;  
 CC Name=CR; Synonyms=BACE-1CR, BACE-1-824;  
 CC Name=CS; Synonyms=BACE-1CS, BACE-1-825;  
 CC Name=CT; Synonyms=BACE-1CT, BACE-1-826;  
 CC Name=CU; Synonyms=BACE-1CU, BACE-1-827;  
 CC Name=CV; Synonyms=BACE-1CV, BACE-1-828;  
 CC Name=CW; Synonyms=BACE-1CW, BACE-1-829;  
 CC Name=CX; Synonyms=BACE-1CX, BACE-1-830;  
 CC Name=CY; Synonyms=BACE-1CY, BACE-1-831;  
 CC Name=CA; Synonyms=BACE-1CA, BACE-1-832;  
 CC Name=CB; Synonyms=BACE-1CB, BACE-1-833;  
 CC Name=CC; Synonyms=BACE-1CC, BACE-1-834;  
 CC Name=CD; Synonyms=BACE-1CD, BACE-1-835;  
 CC Name=CE; Synonyms=BACE-1CE, BACE-1-836;  
 CC Name=CF; Synonyms=BACE-1CF, BACE-1-837;  
 CC Name=CG; Synonyms=BACE-1CG, BACE-1-838;  
 CC Name=CH; Synonyms=BACE-1CH, BACE-1-839;  
 CC Name=CI; Synonyms=BACE-1CI, BACE-1-840;  
 CC Name=CJ; Synonyms=BACE-1CJ, BACE-1-841;  
 CC Name=CK; Synonyms=BACE-1CK, BACE-1-842;  
 CC Name=CL; Synonyms=BACE-1CL, BACE-1-843;  
 CC Name=CM; Synonyms=BACE-1CM, BACE-1-844;  
 CC Name=CN; Synonyms=BACE-1CN, BACE-1-845;  
 CC Name=CO; Synonyms=BACE-1CO, BACE-1-846;  
 CC Name=CP; Synonyms=BACE-1CP, BACE-1-847;  
 CC Name=CQ; Synonyms=BACE-1CQ, BACE-1-848;  
 CC Name=CR; Synonyms=BACE-1CR, BACE-1-849;  
 CC Name=CS; Synonyms=BACE-1CS, BACE-1-850;  
 CC Name=CT; Synonyms=BACE-1CT, BACE-1-851;  
 CC Name=CU; Synonyms=BACE-1CU, BACE-1-852;  
 CC Name=CV; Synonyms=BACE-1CV, BACE-1-853;  
 CC Name=CW; Synonyms=BACE-1CW, BACE-1-854;  
 CC Name=CX; Synonyms=BACE-1CX, BACE-1-855;  
 CC Name=CY; Synonyms=BACE-1CY, B



FT STRAND 211 215  
FT HELIX 224 229

Query Match 99.1%; Score 2267; DB 1; Length 501;  
Best Local Similarity 99.1%; Pred. No. 5.3e-174;  
Matches 429; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```
QY 1 TONGIRLPLRSGLGAPLGLEINLETDEPEEPGRGSGFVEMVDNLRGSGGYVEMTV 60
DB 22 TONGIRLPLRSGLGAPLGRLRPRTDEPEEPGRGSGFVEMVDNLRGSGGYVEMTV 81
QY 61 GSPPTQNLIVDTGSSNFAVGAAPPLHRRYQROLSTYRDLRGVVPVPTQGMGEL 120
DB 82 GSPPTQNLIVDTGSSNFAVGAAPPLHRRYQROLSTYRDLRGVVPVPTQGMGEL 141
QY 121 GTDLVSIHGPVNTVRANIAAITEEDKFFINGSMNEGILGLAYAEIARPDLSLEFPDNL 180
DB 142 GTDLVSIHGPVNTVRANIAAITEEDKFFINGSMNEGILGLAYAEIARPDLSLEFPDNL 201
QY 181 VKQTHVPMFLSIQLCGAGFPPLNQSSEVLASVGGSMITGGIDHSLYTGSLMYTPIRREWYE 240
DB 202 VKQTHVPMFLSIQLCGAGFPPLNQSSEVLASVGGSMITGGIDHSLYTGSLMYTPIRREWYE 261
QY 241 VIVAVEINGODLKMDCCKEYNDKSIIVDSGTTNLRPKKVFEAAVKSIIKAASSTKFPDG 300
DB 262 VIVAVEINGODLKMDCCKEYNDKSIIVDSGTTNLRPKKVFEAAVKSIIKAASSTKFPDG 321
QY 301 FVLGSQLVCWQAGTTPMNIIFPVISLYLMGEVTVNOSFRITILPQOYLRPVEDVATSDQDCY 360
DB 322 FVLGSQLVCWQAGTTPMNIIFPVISLYLMGEVTVNOSFRITILPQOYLRPVEDVATSDQDCY 381
QY 361 KFAISQSTGTGAVMEGFVVDPRARKRIGFAVSAACHVDEPRTAAVEGPFYTLDM 420
DB 382 KFAISQSTGTGAVMEGFVVDPRARKRIGFAVSAACHVDEPRTAAVEGPFYTLDM 441
QY 421 DCGYNIPTDEBT 433
DB 442 DCGYNIPTDEBT 454
```

## RESULT 2

```
Q8IYC8 PRELIMINARY; PRT; 501 AA.
AC Q8IYC8;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DE 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Beta-site App-cleaving enzyme 1, isoform A preproprotein.
GN Name=BACE1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_Taxid=9606;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnae.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schler G.D.,
RA Altshuler S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Mullaney S.J.,
RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullaney S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Guaratine P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Blatterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Scherch A., Schein J.E.,
```

RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences."  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RN (2)  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;

RA Strausberg R.;  
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.

CC -1 SIMILARITY: Belongs to peptidase family A1.

DR EMBL: BC036084; AAH36084.1; -.

DR HSSP: P56817; 1FKN.

DR GO: GO:0005768; C:endosome; ISS.

DR GO: GO:0005794; C:Golgi apparatus; ISS.

DR GO: GO:0016021; C:integral to membrane; ISS.

DR GO: GO:0004190; F:aspartic-type endopeptidase activity; ISS.

DR GO: GO:0050435; P:beta-amyloid metabolism; ISS.

DR GO: GO:0006509; P:membrane protein ectodomain proteolysis; ISS.

DR InterPro: IPR001461; Peptidase A1.

DR InterPro: IPR009119; Pept\_A1\_BACE.

DR InterPro: IPR009120; Pept\_A1\_BACE1.

DR InterPro: IPR009007; Pept\_Aspartic.

DR InterPro: IPR001969; Pept\_Asp\_AS.

DR PRINTS: PRO1816; BACE1.

DR PRINTS: PRO1815; BACEFAMILY.

DR PRINTS: PRO0792; PERSIN.

DR PROSITE: PS00141; ASP\_PROTEASE; 1.

DR Aspartyl1 protease; Hydrolyase; Protease;

SEQUENCE 501 AA; 55823 MW; 768595CF5517EFB7 CRC64;

Query Match 98.8%; Score 2261; DB 2; Length 501;

Best Local Similarity 98.8%; Pred. No. 1.6e-173;  
Matches 428; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```
QY 1 TONGIRLPLRSGLGAPLGLEINLETDEPEEPGRGSGFVEMVDNLRGSGGYVEMTV 60
DB 22 TONGIRLPLRSGLGAPLGRLRPRTDEPEEPGRGSGFVEMVDNLRGSGGYVEMTV 81
QY 61 GSPPTQNLIVDTGSSNFAVGAAPPLHRRYQROLSTYRDLRGVVPVPTQGMGEL 120
DB 82 GSPPTQNLIVDTGSSNFAVGAAPPLHRRYQROLSTYRDLRGVVPVPTQGMGEL 141
QY 121 GTDLVSIHGPVNTVRANIAAITEEDKFFINGSMNEGILGLAYAEIARPDLSLEFPDNL 180
DB 142 GTDLVSIHGPVNTVRANIAAITEEDKFFINGSMNEGILGLAYAEIARPDLSLEFPDNL 201
QY 181 VKQTHVPMFLSIQLCGAGFPPLNQSSEVLASVGGSMITGGIDHSLYTGSLMYTPIRREWYE 240
DB 202 VKQTHVPMFLSIQLCGAGFPPLNQSSEVLASVGGSMITGGIDHSLYTGSLMYTPIRREWYE 261
QY 241 VIVAVEINGODLKMDCCKEYNDKSIIVDSGTTNLRPKKVFEAAVKSIIKAASSTKFPDG 300
DB 262 VIVAVEINGODLKMDCCKEYNDKSIIVDSGTTNLRPKKVFEAAVKSIIKAASSTKFPDG 321
QY 301 FVLGSQLVCWQAGTTPMNIIFPVISLYLMGEVTVNOSFRITILPQOYLRPVEDVATSDQDCY 360
DB 322 FVLGSQLVCWQAGTTPMNIIFPVISLYLMGEVTVNOSFRITILPQOYLRPVEDVATSDQDCY 381
QY 361 KFAISQSTGTGAVMEGFVVDPRARKRIGFAVSAACHVDEPRTAAVEGPFYTLDM 420
DB 382 KFAISQSTGTGAVMEGFVVDPRARKRIGFAVSAACHVDEPRTAAVEGPFYTLDM 441
QY 421 DCGYNIPTDEBT 433
DB 442 DCGYNIPTDEBT 454
```

## RESULT 3

```
BAE1 RAT STANDARD; PRT; 501 AA.
ID BAE1 RAT
AC P56819;
DT 30-MAY-2000 (rel. 39, Last sequence update)
```

DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Beta-secretase 1 precursor (EC 3.4.23.46) (Beta-site APP cleaving  
DE enzyme 1) (Beta-site amyloid precursor protein cleaving enzyme 1)  
DE (Aspartyl) protease 2 (Asp 2) (ASP2) (Membrane-associated aspartic  
DE protease 2) (Memapsin-2).  
GN Name=BACE1; Synonyms=BACE;  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20002972; PubMed=10531052; DOI=10.1126/science.286.5440.735;  
RA Vassar R., Bennett B.D., Babu-Khan S., Kahn S., Mendiaz E.A.,  
RA Denis P., Teplow D.B., Ross S., Amarante P., Loeloff R., Luo Y.,  
RA Fisher S., Fuller J., Edenson S., Lile J., Jarosinski M.A.,  
RA Biere A.L., Curran E., Burgess T., Louis J.-C., Collins F.,  
RA Treanor J., Rogers G., Citron M.,  
RT "Beta-secretase cleavage of Alzheimer's amyloid precursor protein by  
RT the transmembrane aspartic protease BACE.";  
RL Science 286:735-741(1999).  
CC -! FUNCTION: Responsible for the proteolytic processing of the  
CC amyloid precursor protein (APP). Cleaves at the amino terminus of  
CC the A-beta peptide sequence, between residues 671 and 672 of APP,  
CC leads to the generation and extracellular release of beta-cleaved  
CC soluble APP, and a corresponding cell-associated carboxy-terminal  
CC fragment which is later released by gamma-secretase (By  
CC similarity).  
CC -! CATALYTIC ACTIVITY: Broad endopeptidase specificity. Cleaves Glu-  
CC Val-Ala-Leu-Ileu-Ala-Glu-Phe in the Swedish variant of  
CC Alzheimer's amyloid precursor protein.  
CC -! SUBCELLULAR LOCATION: Type I membrane protein.  
CC -! SIMILARITY: Belongs to the peptidase A1 family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AF190727; AAF04144.1; -.  
DR HSSP; P56817; 1M4H.  
DR MEROPS; A01.004; -.  
DR RCD; 2191; Bace.  
DR InterPro; IPR009119; Pept\_A1\_BACE.  
DR InterPro; IPR009120; Pept\_A1\_BACE1.  
DR InterPro; IPR001969; Pept\_Asp\_AS.  
DR InterPro; IPR009007; Pept\_Aspartic.  
DR InterPro; IPR001461; Peptidase\_A1.  
DR Pfam; PF00026; Asp\_1.  
DR PRINTS; PRO1816; BACE1.  
DR PRINTS; PRO1815; BACEFAMILY.  
DR PRINTS; PRO0792; PEPsin.  
DR PROSITE; PS00141; ASP\_PROTEASE; 1.  
DR KMW Aspartyl protease; Glycoprotein; Hydrolyase; Signal; Transmembrane;  
DR Zymogen.  
FT SIGNAL 1 21 Potential.  
FT PROPEP 22 45 Potential.  
FT CHAIN 46 501 Beta-secretase 1.  
FT DOMAIN 22 457 Extracellular (Potential).  
FT TRANSMEM 458 478 Potential.  
FT DOMAIN 479 501 Cytoplasmic (Potential).  
FT ACT\_SITE 93 93 By similarity.  
FT ACT\_SITE 289 289 By similarity.  
FT DISULFID 216 420 By similarity.  
FT DISULFID 278 443 By similarity.  
FT DISULFID 330 380 By similarity.  
FT CARBOHYD 153 153 N-linked (GlcNAc...)(Potential).  
FT CARBOHYD 172 172 N-linked (GlcNAc...)(Potential).  
FT CARBOHYD 223 223 N-linked (GlcNAc...)(Potential).  
FT CARBOHYD 354 354 N-linked (GlcNAc...)(Potential).  
FT

SQ SEQUENCE 501 AA; 55806 MW; 24B45BCBBE87DE3 CRC64;  
Query Match 96.9%; Score 2218; DB 1; Length 501;  
Best Local Similarity 96.5%; Pred. No. 4.7e-170;  
Matches 418; Conservative 5; Mismatches 10; Indels 0; Gaps 0;  
QY 1 TGHGRLPLRSGLGAPILGLEINLETDEEPERGRSGFVEMVNLKSGSGGYVMV 60  
DB 22 THGIRLPLRSGLAPPLGLRLPRITDEEPERGRSGFVEMVNLKSGSGGYVMV 81  
QY 61 GSPPTLILVDYTGSSNFAVGAAPHPPLHRYQQLSTYRDLKSGYVPTGKWEGL 120  
DB 82 GSPPTLILVDYTGSSNFAVGAAPHPPLHRYQQLSTYRDLKSGYVPTGKWEGL 141  
QY 121 GTDLVSIHGNVTVYRANIAITSDKFTNGSWMEIIGLAVALRPPDSLEPPDSL 180  
DB 142 GTDLVSIHGNVTVYRANIAITSDKFTNGSWMEIIGLAVALRPPDSLEPPDSL 201  
QY 161 VKQTHVPMPLSLQCGAGFPLNGSEVLAVSGSMITGIDHSLYTGSLMYTPIRREMYE 240  
DB 202 VKQTHVPMPLSLQCGAGFPLNGSEVLAVSGSMITGIDHSLYTGSLMYTPIRREMYE 261  
QY 241 VIVRVEINQDLMKDKENYVDKSYDSTGTMRLPKYFEAAVKSIRKASTETKPPDG 300  
DB 262 VIVRVEINQDLMKDKENYVDKSYDSTGTMRLPKYFEAAVKSIRKASTETKPPDG 321  
QY 301 FWLGEQLVCWQAGTTPNNIFPVISLYLGEVTNQSFRTITLPQOYLRPEVDVTSQDDCY 360  
DB 322 FWLGEQLVCWQAGTTPNNIFPVISLYLGEVTNQSFRTITLPQOYLRPEVDVTSQDDCY 381  
QY 361 KEAISQSTGTWCAVIMGEFYVFPDARKRIGFAVSACHVDEFTAAVEGPPVTLME 420  
DB 382 KEAISQSTGTWCAVIMGEFYVFPDARKRIGFAVSACHVDEFTAAVEGPPVTLME 441  
QY 421 DCGYNIPQTDST 433  
DB 442 DCGYNIPQTDST 454  
RESULT 4  
BAE1 MOUSE STANDARD; PRT; 501 AA.  
AC P56818;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Beta-secretase 1 precursor (EC 3.4.23.46) (Beta-site APP cleaving  
DE enzyme 1) (Beta-site amyloid precursor protein cleaving enzyme 1)  
DE (Aspartyl) protease 2 (Asp 2) (ASP2) (Membrane-associated aspartic  
DE protease 2) (Memapsin-2).  
GN Name=BACE1; Synonyms=BACE;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20002972; PubMed=10531052; DOI=10.1126/science.286.5440.735;  
RA Vassar R., Bennett B.D., Babu-Khan S., Kahn S., Mendiaz E.A.,  
RA Denis P., Teplow D.B., Ross S., Amarante P., Loeloff R., Luo Y.,  
RA Fisher S., Fuller J., Edenson S., Lile J., Jarosinski M.A.,  
RA Biere A.L., Curran E., Burgess T., Louis J.-C., Collins F.,  
RA Treanor J., Rogers G., Citron M.,  
RT "Beta-secretase cleavage of Alzheimer's amyloid precursor protein by  
RT the transmembrane aspartic protease BACE.";  
RL Science 286:735-741(1999).  
RN [2]  
RP REVISIONS TO 6 AND 81-87.  
RA Bennett B.D., Vassar R., Citron M.,  
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20057170; PubMed=10591213; DOI=10.1038/990107;

RA van R., Binkowski M.J., Shuck M.E., Miao H., Torry M.C., Pauley A.M.,  
 RA Braahler J.R., Stratman N.C., Mathew W.R., Buhl A.E., Carter D.B.,  
 RA Tomasselli A.G., Parodi L.A., Heinrikson R.L., Gurney M.E.,  
 RT "Membrane-anchored aspartyl protease with Alzheimer's disease beta-  
 RT secretase activity." ;  
 RL Nature 402:533-537(1999).  
 (4)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Brain;  
 RX MEDLINE=22354663; PubMed=12466851; DOI=10.1038/nature01266;  
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
 RA Nikaido I., Oosato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,  
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gotohori T.,  
 RA Balderelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,  
 RA Blake J.A., Brad D., Brusic V., Chottha C., Corbani L.E., Cousins S.,  
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Fraser K.S.,  
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,  
 RA Grimond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,  
 RA Kongaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
 RA Maglott D.R., Maltsev L., Marchionni L., McKenzie L., Miki H.,  
 RA Nagashima T., Numa K., Okido T., Pavan W.J., Pereira G., Pesole G.,  
 RA Petrovsky N., Pillai R., Pontius U.U., Qi D., Ramachandran S.,  
 RA Raveisi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
 RA Sandelin A., Schneider C., Sempke C.A., Setou M., Shimada M.,  
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,  
 RA Wilming L.G., Wymshew-Boris A., Yanagisawa M., Yang I., Yang L.,  
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
 RA Hirozane-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,  
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arikawa T., Fukuda S.,  
 RA Hara A., Hashikume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,  
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
 RA Birney E., Hayashizaki Y.,  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs." ;  
 RL Nature 420:563-573(2002).  
 (5)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Brain;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heile F.,  
 RA Diatchenko L., Marusik A., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loughellano N.A., Peters G.J., Abramson R.D., Muliyil S.J.,  
 RA Bosak S.S., McKean P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hylk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.W., Marra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences." ;  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 -1- FUNCTION: Responsible for the proteolytic processing of the  
 amyloid precursor protein (APP). Cleaves at the amino termini of  
 the A-beta peptide sequence, between residues 671 and 672 of APP,  
 leads to the generation and extracellular release of beta-cleaved  
 soluble APP, and a corresponding cell-associated carboxy-terminal  
 fragment which is later released by gamma-secretase (By  
 similarity).  
 -1- CATALYTIC ACTIVITY: Broad endopeptidase specificity. Cleaves Glu-  
 Val-Leu-Ile-Asp-Ala-Glu-Phe in the Swedish variant of  
 Alzheimer's amyloid precursor protein.

CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- TISSUE SPECIFICITY: Brain.  
 CC -1- SIMILARITY: Belongs to the peptidase A1 family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sdb.ch/announce/>  
 CC or send an email to [license@sdb.ch](mailto:license@sdb.ch)).  
 CC -----  
 CC EMBL, AF190726; AAF04143.2; -  
 CC EMBL, AF200346; AAF17082.1; -  
 CC EMBL, AK014464; BAB29370.1; -  
 CC EMBL, BC048189; AAB48189.1; -  
 CC HSSP; P56817; 1M4H.  
 CC MEROPS; A01.004; -  
 CC MGD; MGI:1346542; Bace1.  
 CC InterPro; IPR009119; Pept\_A1\_BACE.  
 CC InterPro; IPR009120; Pept\_A1\_BACE1.  
 CC InterPro; IPR001969; Pept\_Asp\_AS.  
 CC InterPro; IPR009007; Pept\_AspArtic.  
 CC InterPro; IPR001461; Peptidase\_A1.  
 CC Pfam; PF00026; Asp; 1.  
 CC PRINTS; PR01816; BACE1.  
 CC PRINTS; PR01815; BACEFAMILY.  
 CC PRINTS; PR00792; PEPsin.  
 CC PROSITE; PS00141; Asp\_PROTEASE; 1.  
 CC Aspartyl protease; Glycoprotein; Hydrolase; Signal; Transmembrane;  
 KW Zymogen.  
 FT SIGNAL 1 21 Potential.  
 FT PROPEP 22 45 Potential.  
 FT CHAIN 46 501 Beta-secretase 1.  
 FT DOMAIN 22 457 Extracellular (Potential).  
 FT TRANSMEM 458 478 Potential.  
 FT DOMAIN 479 501 Cytoplasmic (Potential).  
 FT ACT\_SITE 93 93 By similarity.  
 FT ACT\_SITE 289 289 By similarity.  
 FT DISULFID 216 420 By similarity.  
 FT DISULFID 278 443 By similarity.  
 FT DISULFID 330 380 By similarity.  
 FT CARBOHYD 153 153 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 172 172 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 223 223 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 354 354 N-linked (GlcNAc...) (Potential).  
 FT SEQUENCE 501 AA; 55747 MW; C085A013145B474E CRC64;  
 Query Match 96.9%; Score 2216; DB 1; Length 501;  
 Best Local Similarity 96.5%; Pred. No. 6; 9e-170;  
 Matches 418; Conservative 5; Mismatches 10; Indels 0; Gaps 0;  
 Oy 1 THGIRLPRLSGAGPLGLGELINLTDEBEPGRGRGSFVENVNLRGSGGYVEMTV 60  
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 Db 22 THIGIRLPRLSGAGPLGLRLPRETDESEBGRGRGSFVENVNLRGSGGYVEMTV 81  
 Oy 61 GSPPOGLNLTIVDTGSSNFAVGAAPHPFLRRYQROQLSSTRDLRKQVYPTYGKKEGL 120  
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 Db 82 GSPPOGLNLTIVDTGSSNFAVGAAPHPFLRRYQROQLSSTRDLRKQVYPTYGKKEGL 141  
 Oy 121 GPDVLSIPHGPNVTPYANIAITESPGRFINSNMGGIGLVAETARPDSDLEPFDSL 180  
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 Db 142 GPDVLSIPHGPNVTPYANIAITESPGRFINSNMGGIGLVAETARPDSDLEPFDSL 201  
 Oy 181 VQOTHPNLFSLQCGAGPFLNQSEVTLASVGSMTIGGIDHSLYTGSLWYTPIRREWYE 240  
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 Db 202 VQOTHPNLFSLQCGAGPFLNQSEVTLASVGSMTIGGIDHSLYTGSLWYTPIRREWYE 261  
 Oy 241 VIVRVEINGQDLKMDCKEYNTDKSVDSGTTNLRPKVYFPAVAVKSTQASSTKFPDG 300  
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 Db 262 VIVRVEINGQDLKMDCKEYNTDKSVDSGTTNLRPKVYFPAVAVKSTQASSTKFPDG 321  
 Oy 301 FVLGSDLVCMQAGTTFWNIFPVISLYLWGEVTVNQSRITLPPQVYLRVEDVATSDDCY 360

Db 322 FWLGEOLVCMQAGTTPMNI.FPVISLYLWGEVNTOSFRITLLPQOYLREVEVATSQDDCY 381  
Qy 361 KRAISQSGSTGYMGAVIMEGYYVFPDRARRKIGFAVSACHVDEFRITAAVEGPFVTLDME 420  
Db 382 KRAVSQSSSTGYMGAVIMEGYYVFPDRARRKIGFAVSACHVDEFRITAAVEGPFVTLDME 441  
Qy 421 DCGYNIPTDDEST 433  
Db 442 DCGYNIPTDDEST 454  
RESULT 5  
Q8C7R1 PRELIMINARY: PRT: 501 AA.  
AC Q8C7R1;  
DT 01-MAR-2003 (TREMBLrel. 23, Created)  
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
DE Mus musculus 12 days embryo spinal cord cDNA, RIKEN full-length  
DE enriched library, clone: C530008K17 product: beta-site APP cleaving  
DE enzyme, full insert sequence.  
GN Name=Bace1; Synonym=Bace1;  
OS Mus musculus (Mouse).  
OC Eukaryota; Euteleostomi; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Spinal cord;  
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;  
RA Carninci P., Hayashizaki Y.;  
RT "High-efficiency full-length cDNA cloning.";  
RL Mech. Enzymol. 303:19-44(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Spinal cord;  
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;  
RA RIKEN FANTOM Consortium;  
RT "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690(2001).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Spinal cord;  
RA The FANTOM Consortium;  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs.";  
RL Nature 420:563-573(2002).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Spinal cord;  
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;  
RA Carninci P., Shibata Y., Hayatsu M., Sugahara Y., Shibata K., Itoh M.,  
RA Komori H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
RT prepare full-length cDNA libraries for rapid discovery of new genes.";  
RL Genome Res. 10:1617-1630(2000).  
RN [5]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Spinal cord;  
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,  
RA Komori H., Akiyama J., Nishi K., Katsunai T., Tashiro H., Itoh M.,  
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kaishima K.,  
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Matsubara S.,  
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,  
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
RT "RIKEN integrated sequence analysis (RISA) system-384-format  
RT sequencing pipeline with 384 multicapillary sequencer.";  
RL Genome Res. 10:1757-1771(2000).  
RN [6]

RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Spinal cord;  
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,  
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,  
RA Hayashida K., Hayatsu M., Hiramoto K., Hirooka T., Hirozane T.,  
RA Horii F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,  
RA Kato H., Kawai J., Kojima Y., Kondo S., Komori H., Koyama S.,  
RA Kuribara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,  
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,  
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,  
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,  
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,  
RA Tomaru A., Toya T., Yasumishi A., Muramatsu M., Hayashizaki Y.,  
RL Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.  
CC -1- SIMILARITY: Belongs to peptidase family A1.  
DR EMBL: AK049626; BAC3844.1; .  
DR HSSP: p56817; 1PRN.  
DR MCD: MCI:1346542; Bace1.  
DR GO: GO:0005768; C:endosome; ISS.  
DR GO: GO:0005615; C:extracellular space; TAS.  
DR GO: GO:0005794; C:Golgi apparatus; ISS.  
DR GO: GO:0016021; C:integral to membrane; ISS.  
DR GO: GO:0004190; F:aspartic-type endopeptidase activity; ISS.  
DR GO: GO:0050435; P:beta-amyloid metabolism; ISS.  
DR GO: GO:0006509; P:membrane protein ectodomain proteolysis; ISS.  
DR InterPro: IPR001461; Peptidase A1.  
DR InterPro: IPR009119; Pept\_A1\_BACE1.  
DR InterPro: IPR009120; Pept\_A1\_BACE1.  
DR InterPro: IPR009007; Pept\_Aspartic.  
DR PRINTS: PRO1816; Pept\_Asp\_AS.  
DR PRINTS: PRO1815; BACE1.  
DR PRINTS: PRO0792; PEPsin.  
DR PROSITE: PS00141; ASP\_PROTEASE; 1.  
KW Aspartyl protease; Hydrolase; Protease.  
SQ SEQUENCE 501 AA; 55761 MW; B410DAB64647663 CRC64;  
Query Match 96.8%; Score 2215; DB 2; Length 501;  
Best Local Similarity 96.3%; Pred. No. 8.3e-170;  
Matches 417; Conservative 6; Mismatches 10; Indels 0; Gaps 0;  
Qy 1 TOHGIRLPLRSGLGAPLGLIEINLFTDEPEERGRGSPVENVMDLRGSGGQYVEMTV 60  
Db 22 THGIRLPLRSGLAGPLGLRLPRETDESEEPGRGSPVENVMDLRGSGGQYVEMTV 81  
Qy 61 GSPPTLNLIVDTGSSNFAVGAAPHFLLHRYYQRLSSTYRDLRGVYVPTGKMEGEL 120  
Db 82 GSPPTLNLIVDTGSSNFAVGAAPHFLLHRYYQRLSSTYRDLRGVYVPTGKMEGEL 141  
Qy 121 GTPLVSIPIHGPNTVYRANIAITSDKPFINSNNEGILGLAYAIAPDDSLPEFPDSL 180  
Db 142 GTPLVSIPIHGPNTVYRANIAITSDKPFINSNNEGILGLAYAIAPDDSLPEFPDSL 201  
Qy 181 VQOTHPNLFSLQCGAPPLNQSEVLASVGSMTIGIDISLVTGSLMTYPIRREYWE 240  
Db 202 VQOTHPNLFSLQCGAPPLNQSEVLASVGSMTIGIDISLVTGSLMTYPIRREYWE 261  
Qy 241 VTIIVREINGDPLKMDCKEYNDKSIYDSGTTNLRPKVFEAAVKSIKAASSTKPPDG 300  
Db 262 VTIIVREINGDPLKMDCKEYNDKSIYDSGTTNLRPKVFEAAVKSIKAASSTKPPDG 321  
Qy 301 FWLGEOLVCMQAGTTPMNI.FPVISLYLWGEVNTOSFRITLLPQOYLREVEVATSQDDCY 360  
Db 322 FWLGEOLVCMQAGTTPMNI.FPVISLYLWGEVNTOSFRITLLPQOYLREVEVATSQDDCY 381  
Qy 361 KRAISQSGSTGYMGAVIMEGYYVFPDRARRKIGFAVSACHVDEFRITAAVEGPFVTLDME 420  
Db 382 KRAVSQSSSTGYMGAVIMEGYYVFPDRARRKIGFAVSACHVDEFRITAAVEGPFVTLDME 441  
Qy 421 DCGYNIPTDDEST 433  
Db 442 DCGYNIPTDDEST 454

RESULT 6  
ID Q8BOY4 PRELIMINARY; PRT; 501 AA.  
AC Q8BOY4;  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Mus musculus adult male corpora quadrigemina cDNA, RIKEN full-length  
DE enriched library, clone:8230346M13 product:beta-site APP cleaving  
DE enzyme, full insert sequence.  
CN Name=Bacel; Synonyms=Bace;  
OS Mus musculus (Mouse);  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_Taxid=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;  
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;  
RA Carinci P., Hayashizaki Y.;  
RT "High-efficiency full-length cDNA cloning."; Meth. Enzymol. 303:19-44(1999).  
RL Nature 409:685-690(2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;  
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;  
RA RIKEN PANTOM Consortium;  
RT "Functional annotation of a full-length mouse cDNA collection."; Nature 409:685-690(2001).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;  
RX MEDLINE=2049374; PubMed=11042159; DOI=10.1101/gr.145100;  
RA Carinci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
RT prepare full-length cDNA libraries for rapid discovery of new genes."; Genome Res. 10:1617-1630(2000).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;  
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carinci P.,  
RA Kono H., Akiyama J., Nishi K., Kitsuina T., Tashiro H., Itoh M.,  
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Wataniki M.,  
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsumura S., Kawai J.,  
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
RT "RIKEN integrated sequence analysis (RISA) system-384-format  
RT sequencing pipeline with 384 multicapillary sequencer."; Genome Res. 10:1757-1771(2000).  
RN [5]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;  
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carinci P.,  
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,  
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,  
RA Horii F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasubawa T.,  
RA Kato H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,  
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,  
RA Nishi K., Nomura K., Numasaki R., Ohno M., Ohnato N., Okazaki Y.,  
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,  
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,

RA Tagawa A., Takahashi F., Takaku-Akaiura S., Takeda Y., Tanaka T.,  
RA Tomaru A., Toyota T., Yasunishi A., Muramatsu M., Hayashizaki Y.;  
RL Submitted (Jul-2001) to the EMBL/Genbank/DBJ databases.  
CC -1- SIMILARITY: Belongs to peptidase family A1.  
DR EMBL; AK046175; BAC32620.1; -.  
DR HSSP; P56817; 1FKN.  
DR MGD; MGI:1346542; Bace1.  
DR GO; GO:0005768; C:endosome; ISS.  
DR GO; GO:0005615; C:extracellular space; TAS.  
DR GO; GO:0005794; C:oligodendrocyte; ISS.  
DR GO; GO:0016021; C:integral to membrane; ISS.  
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; ISS.  
DR GO; GO:0050435; P:beta-amyloid metabolism; ISS.  
DR GO; GO:0005509; P:membrane protein ectodomain proteolysis; ISS.  
DR InterPro; IPR001461; Peptidase\_A1.  
DR InterPro; IPR009119; Pept\_A1\_BACE.  
DR InterPro; IPR009120; Pept\_A1\_BACE.  
DR InterPro; IPR009007; Pept\_Aspartic.  
DR InterPro; IPR01969; Pept\_Asp\_AS.  
DR PRINTS; PR01816; BACE1.  
DR PRINTS; PR01815; BACEFAMILY.  
DR PRINTS; PR00792; PEPSTN.  
DR PROSITE; PS00141; ASP\_PROTEASE; 1.  
KW Aspartyl protease; Hydrolase; Protease.  
SQ SEQUENCE 501 AA; 55816 MW; C085513145B024E CRC64;  
Query Match 96.6%; Score 2211; DB 2; Length 501;  
Best Local Similarity 96.3%; Pred. No. 1,7e-165;  
Matches 417; Conservative 5; Mismatches 11; Indels 0; Gaps 0;  
QY 1 TQHGIRLPFRSGLGAPLGLEINLETDEPEPEGRGSGFVEMVDNLRKSGGGYVEMTV 60  
DB 22 THLGIRLPFRSGLGAPLGLEINLETDEPEPEGRGSGFVEMVDNLRKSGGGYVEMTV 81  
QY 61 GSPPTQNLIVDTGSSNFAVGAAPFLHRYRQLSSTYRDLRGVYVPTQGWEGEL 120  
DB 82 GSPPTQNLIVDTGSSNFAVGAAPFLHRYRQLSSTYRDLRGVYVPTQGWEGEL 141  
QY 121 GTDLVSIHPGPNVTVRANIAATTESDKFFINSNMEGILGLAYATARPDDSLPEPFDL 180  
DB 142 GTDLVSIHPGPNVTVRANIAATTESDKFFINSNMEGILGLAYATARPDDSLPEPFDL 201  
QY 181 VQOTHPNMFSLQCGAGPPLNOSVLASVGSMTIIGIDHSLYTGSMTYTPRREWEYE 240  
DB 202 VQOTHPNMFSLQCGAGPPLNOSVLASVGSMTIIGIDHSLYTGSMTYTPRREWEYE 261  
QY 241 VTIIVEINGQDLKMDCKEYNDKSIIVDSGTNLRPKVFBAVKSIAASTEKFPDG 300  
DB 262 VTIIVEINGQDLKMDCKEYNDKSIIVDSGTNLRPKVFBAVKSIAASTEKFPDG 321  
QY 301 FWLGEOLVQWAGTTPWNIFFVISTYLMGEVTNOSFRITILPQOYLRPVEDVATGDDCY 360  
DB 322 FWLGEOLVQWAGTTPWNIFFVISTYLMGEVTNOSFRITILPQOYLRPVEDVATGDDCY 381  
QY 361 KRAISGSSTGTMGAVIMEGFVVPDRARKRIGFVPSACHVDEFTAAVEGPFYTLDMWE 420  
DB 382 KRAVSSSTGTMGAVIMEGFVVPDRARKRIGFVPSACHVDEFTAAVEGPFYTLDMWE 441  
QY 421 DCGYNIPTDDEST 433  
DB 442 DCGYNIPTDDEST 454  
RESULT 7  
ID Q9UJ51 PRELIMINARY; PRT; 532 AA.  
AC Q9UJ51;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE KIAA1149 protein (Fragment).  
GN Name=KIAA1149;  
OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
NCBI\_TaxId=9606;  
RN (1)  
RC SEQUENCE FROM N.A.  
RX MEDLINE=20039618; PubMed=10574461;  
RA Hirogawa M., Nagase T., Ishikawa K., Kikuno R., Nomura N., Ohara O.,  
RT "Characterization of cDNA clones selected by the Genemark analysis  
RL from size-fractionated cDNA libraries from human brain.";  
CC DNA Res. 6:329-336(1999).  
DR -1- SIMILARITY: Belongs to peptidase family A1.  
DR EMBL; AB032975; BAA86463.2; -.  
DR HSSP; P56817; 1PKN.  
DR GO; GO:0005768; C:endosome; ISS.  
DR GO; GO:0005794; C:Golgi apparatus; ISS.  
DR GO; GO:0016021; C:integral to membrane; ISS.  
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; ISS.  
DR GO; GO:0050435; P:beta-amyloid metabolism; ISS.  
DR GO; GO:0006509; P:membrane protein ectodomain proteolysis; ISS.  
DR InterPro; IPR001461; Peptidase A1.  
DR InterPro; IPR009119; Pept\_A1\_BACE.  
DR InterPro; IPR009120; Pept\_A1\_BACE1.  
DR InterPro; IPR009007; Pept\_AspArtic.  
DR InterPro; IPR001969; Pept\_Asp\_AS.  
DR PRINTS; PRO1816; BACE1.  
DR PRINTS; PRO1815; BACEFAMILY.  
DR PRINTS; PRO0792; PEPsin.  
DR PROSITE; PS00141; ASP\_PROTEASE; 1.  
DR APPARTYL; protease; Hydrolyase; Protease.  
FT NON\_TER 1  
SQ SEQUENCE 532 AA; 58720 MW; 98B135D05FBD28 CRC64;  
Query Match 96.3%; Score 2204; DB 2; Length 532;  
Best Local Similarity 97.9%; Pred. No. 6.9e-169;  
Matches 417; Conservative 2; Mismatches 7; Indels 0; Gaps 0;  
QY 8 PLASGLGADPLGLEINLETDEEPEEPEGRSGFVEMVDNLRGSGGQYVEMTVGSPPTL 67  
DB 60 PCAAAMGCAPLGLRLPRETDEEPEEPEGRSGFVEMVDNLRGSGGQYVEMTVGSPPTL 119  
QY 68 NILVDGSSNFAVGAAPHPLHRYRORLSTTRDRLKRVYVPTYGKMEGLGDLVSI 127  
DB 120 NILVDGSSNFAVGAAPHPLHRYRORLSTTRDRLKRVYVPTYGKMEGLGDLVSI 179  
QY 128 PHGPNTVANTIAITESDKFINSNMWGIIGLVAEIAFPDDSLPEPDSLVQCHVP 187  
DB 180 PHGPNTVANTIAITESDKFINSNMWGIIGLVAEIAFPDDSLPEPDSLVQCHVP 239  
QY 188 NILFSLQLCGAGPPLNOSVLAIVGSGMITGGIDHSLVYTGSLWYTPIRREWYEVIIYVE 247  
DB 240 NILFSLQLCGAGPPLNOSVLAIVGSGMITGGIDHSLVYTGSLWYTPIRREWYEVIIYVE 299  
QY 248 INGODLMDCKEYNDKSIYDSGTTNLRPKVFEAVAKSIAASTKFPDGLGQL 307  
DB 300 INGODLMDCKEYNDKSIYDSGTTNLRPKVFEAVAKSIAASTKFPDGLGQL 359  
QY 308 VCMQAGTTPWNI FPIYSILYMGVETNOSFRITLPEQYLRPVEDVATSODDCYFAISQS 367  
DB 360 VCMQAGTTPWNI FPIYSILYMGVETNOSFRITLPEQYLRPVEDVATSODDCYFAISQS 419  
QY 368 STGTVMGAVIMGFFVVDRAARKRGFAVSACHVDEERTAAVGPPTLDMCCGNIP 427  
DB 420 STGTVMGAVIMGFFVVDRAARKRGFAVSACHVDEERTAAVGPPTLDMCCGNIP 479  
QY 428 QTDST 433  
DB 480 QTDST 485

AC O8CAF4;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Mus musculus 0 day neonate cerebellum cDNA, RIKEN full-length enriched  
DE library, clone: C30037B16 product: beta-site APP cleaving enzyme, full  
DE insert sequence.  
GN Name=Bace1; Synonyms=Bace;  
OS Mus musculus (Mouse)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NCBI\_TaxId=10090;  
RN (1)  
RC SEQUENCE FROM N.A.  
RX MEDLINE=C57BL/6J; TISSUE=Cerebellum;  
RA Carninci P., Hayashizaki Y.;  
RT "High-efficiency full-length cDNA cloning.";  
RL Meth. Enzymol. 303:19-44(1999).  
RN (2)  
RC SEQUENCE FROM N.A.  
RX STRAIN=C57BL/6J; TISSUE=Cerebellum;  
RA MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;  
RA RIKEN PANTOM Consortium;  
RT "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690(2001).  
RN (3)  
RC SEQUENCE FROM N.A.  
RX STRAIN=C57BL/6J; TISSUE=Cerebellum;  
RA The FANTOM Consortium;  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs.";  
RL Nature 420:563-573(2002).  
RN (4)  
RC SEQUENCE FROM N.A.  
RX STRAIN=C57BL/6J; TISSUE=Cerebellum;  
RA MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;  
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
RT prepare full-length cDNA libraries for rapid discovery of new genes.";  
RL Genome Res. 10:1617-1630(2000).  
RN (5)  
RC SEQUENCE FROM N.A.  
RX STRAIN=C57BL/6J; TISSUE=Cerebellum;  
RA MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,  
RA Kono H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,  
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Washiki M.,  
RA Fujisake S., Inoue K., Togawa K., Ozawa K., Tanaka T., Matsura S., Kawai J.,  
RA Yoneda Y., Ishikawa T., Ozawa K., Izawa M., Ohara E., Kashiwagi K.,  
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
RT "RIKEN integrated sequence analysis (RISA) system-384-format  
RT sequencing pipeline with 384 multicapillary sequencer.";  
RL Genome Res. 10:1757-1771(2000).  
RN (6)  
RC SEQUENCE FROM N.A.  
RX STRAIN=C57BL/6J; TISSUE=Cerebellum;  
RA Adachi U., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,  
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,  
RA Hayashida K., Hayatsu N., Hiramoto K., Hirooka T., Hirozane T.,  
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,  
RA Katoh H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,  
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,  
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohata N., Okazaki Y.,  
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,  
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,  
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,  
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;  
RT Submitted (Apr-2002) to the EMBL/Genbank/DBJ databases.  
CC -1- SIMILARITY: Belongs to peptidase family A1.

DR EMBL; AK082317; BAC38462.1; --  
 DR HSSP; P56817; 1FRN.  
 DR MGD; MGI:1346542; Bace1.  
 DR GO; GO:0005768; C-endosome; ISS.  
 DR GO; GO:0005615; Cextracellular space; TMS.  
 DR GO; GO:0005794; Cgolgi apparatus; ISS.  
 DR GO; GO:0016021; Cintegral to membrane; ISS.  
 DR GO; GO:0004190; F:aspartic-type endopeptidase activity; ISS.  
 DR GO; GO:0005435; P:beta-amyloid metabolism; ISS.  
 DR GO; GO:0005509; P:membrane protein ectodomain proteolysis; ISS.  
 DR InterPro; IPR001461; Peptidase A1.  
 DR InterPro; IPR009119; Pept\_A1\_BACE.  
 DR InterPro; IPR009120; Pept\_A1\_BACE.  
 DR InterPro; IPR009007; Pept\_Aspartic.  
 DR InterPro; IPR001969; Pept\_Asp\_AS.  
 DR PRINTS; PRO1816; BACE1.  
 DR PRINTS; PRO1815; BACEFAMILY.  
 DR PRINTS; PRO0792; PEPsin.  
 DR PROSITE; PS00141; ASP\_PROTEASE; 1.  
 DR Aspartyl protease; Hydrolyase; Protease.  
 KW Aspartyl protease; Hydrolyase; Protease.  
 SQ SEQUENCE 467 AA; 52063 MW; 31AB674FP1843652 CRC64;

Query Match 88.4%; Score 2023; DB 2; Length 467;  
 Best Local Similarity 88.7%; Pred. No. 2,2e-154; Indels 34; Gaps 1;  
 Matches 384; Conservative 5; Mismatches 10;

QY 1 TGHGIRLPARGSLGAPLGLINLETDEPEREPRGRSGFVENVNLRKSGGGYVEMTV 60  
 DB 22 TLHGIRLPARGSLGAPLGLRLPRIDESEBERGRSGFVENVNLRKSGGGYVEMTV 81  
 QY 61 GSPPTQNLIVDTGSSNFAVGAAPHPFLHRYRQRLSSTYRDLRKGVYVPTQKKEGL 120  
 DB 82 GSPPTQNLIVDTGSSNFAVGAAPHPFLHRYRQRLSSTYRDLRKGVYVPTQKKEGL 141  
 QY 121 GRDIVSIPHGPNVTYRANIAATESKPFINSNMGEGILGLAYAEIARPDDELPEPDL 180  
 DB 142 GRDLVSIHPGPNVTYRANIAATESKPFINSNMGEGILGLAYAEIARPDDELPEPDL 201  
 QY 181 VQKTHVFNLFSLQLCAGFPLNQLSEVLAASVSGSMITGIDHSLYTSLWTPPIRMYE 240  
 DB 202 VQKTHVFNLFSLQLCAGFPLNQLSEVLAASVSGSMITGIDHSLYTSLWTPPIRMYE 261  
 QY 241 VTIIVVEINGODLKMDCKEYNDKSIIVDSGTTNLRPKVFEAAVKSIRAASTKPPG 300  
 DB 262 VTIIVVEINGODLKMDCKEYNDKSIIVDSGTTNLRPKVFEAAVKSIRAASTKPPG 287  
 QY 301 FNLGEBLVCMQAGTTTWNIPVISTLYMGEVYTNQSFRTILPQOYLRPVEDVATSDDCY 360  
 DB 288 FNLGEBLVCMQAGTTTWNIPVISTLYMGEVYTNQSFRTILPQOYLRPVEDVATSDDCY 347  
 QY 361 KFAISOSTGTWGAIVMEGFYVFPDRAKRIIGFAVSACHVDEFTAAVEGFTVLDME 420  
 DB 348 KFAVSOSTGTWGAIVMEGFYVFPDRAKRIIGFAVSACHVDEFTAAVEGFTVLDME 407  
 QY 421 DCGYNIPTDEST 433  
 DB 408 DCGYNIPTDEST 420  
 RESULT 9  
 ID Q6NZT7 PRELIMINARY; PRT; 505 AA.  
 AC Q6NZT7;  
 DT 05-JUL-2004 (T-EMBLrel. 27, Created)  
 DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)  
 DE Hypothetical protein zgc:77409.  
 GN Name:zgc:77409;  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OX NCBI\_TaxId=7955;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=embryo;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.J.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McKernan P.J., McKernan K.J., Malek J.A., Guarnarata P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,  
 RA Villalón D.K., Kuzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton B., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smalios D.E., Scherch A., Schein J.E.,  
 RA Jones S.J., Maiz M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RL and mouse cDNA sequences."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]

RP SEQUENCE FROM N.A.  
 RC TISSUE=embryo;  
 RC Director MGC Project;  
 RL Submitted (FEB-2004) to the EMBL/Genbank/DBJ databases.  
 CC -1- SIMILARITY: Belongs to peptidase family A1.  
 DR EMBL; BC065973; AAH65973.1; --  
 DR HSSP; P56272; 1AMS.  
 DR ZFIN; ZDB-GENE-040426-1835; zgc:77409.  
 DR GO; GO:0009049; F:aspartic-type signal peptidase activity; IEA.  
 DR GO; GO:0004194; F:pepsin A activity; IEA.  
 DR GO; GO:0008233; F:peptidase activity; IEA.  
 DR GO; GO:0005508; P:proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR001461; Peptidase A1.  
 DR InterPro; IPR009119; Pept\_A1\_BACE.  
 DR InterPro; IPR009120; Pept\_A1\_BACE.  
 DR InterPro; IPR009007; Pept\_Aspartic.  
 DR InterPro; IPR001969; Pept\_Asp\_AS.  
 DR PRINTS; PRO1816; BACE1.  
 DR PRINTS; PRO1815; BACEFAMILY.  
 DR PRINTS; PRO0792; PEPsin.  
 DR PROSITE; PS00141; ASP\_PROTEASE; 1.  
 KW Aspartyl protease; Hydrolyase; Protease.  
 SQ SEQUENCE 505 AA; 55661 MW; 6C83DA4CFD903BAD CRC64;

Query Match 78.4%; Score 1794; DB 2; Length 505;  
 Best Local Similarity 77.3%; Pred. No. 7.1e-136; Indels 12; Gaps 3;  
 Matches 338; Conservative 40; Mismatches 47;

QY 5 IRLPLRSGLGAPLGLINLETDEPEREPRGRG-----SFENVNLRKSGGGYVE 57  
 DB 28 LRVPLRQG---PLPHSTPAHRSBPHGRASRAASGISTFINMDINLRKSGGGYVE 83  
 QY 58 MTGSPPTQNLIVDTGSSNFAVGAAPHPFLHRYRQRLSSTYRDLRKGVYVPTQKKE 117  
 DB 84 MAVGSPAQRLNIVDTGSSNFAVGAAPHPFLHRYRSLSSYRDLGRGVYVPTQGRWE 143  
 QY 118 GELGDLVSIHPGPNVTYRANIAATESKPFINSNMGEGILGLAYAEIARPDDELPEP 177  
 DB 144 GELGDLVSVHPGPNVTYRANIAATESKPFINSNMGEGILGLAYAEIARPDDELPEP 203  
 QY 178 DSLVQKTHVFNLFSLQLCAGFPLNQLSEVLAASVSGSMITGIDHSLYTSLWTPPIR 236  
 DB 204 DSLRQSTYADVFSIQLCAGFPLNQLSEVLAASVSGSMITGIDHSLYTSLWTPPIR 263  
 QY 237 WYVEVIVVEINGODLKMDCKEYNDKSIIVDSGTTNLRPKVFEAAVKSIRAASTK 296  
 DB 264 WYVEVIVVEINGODLNMDCKEYNDKSIIVDSGTTNLRPKVFEAAVKSIRAASTK 323



QY 297 FPDGFWLGEOLVCMOAGTTPMNIPIVLSLIMGEVTVNSGFRITTLIPQOYLREVEDVATSO 356  
DB 324 FBSGFWLGEOLVCMOAGTTPMNIPIVLSLIMSENNOSFRISITLIPQOYLREVEDVATSO 383  
QY 357 DDCYKRAISQSSGTGTGAVIMEGFYVFPDRAKRGFAVSAACHVDEPRTAAVGEPPVT 416  
DB 384 EDCYKRAVSSGTGTGAVIMEGFYVFEROHKRGFAVSTCHVDEPRTAAVGEPPHG 443  
QY 417 LDMEDCGNINIPQDEST 433  
DB 444 LDEDCGNVNPQDEST 460

## RESULT 10

Q99206 PRELIMINARY; PRT; 373 AA.  
AC 069206: 25-OCT-2004 (TREMBlrel. 28, Created)  
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)  
DE MKIA1149 protein (Fragment).  
GN Name=MKIA1149;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCB1\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryonic tail;  
RA Okazaki N., Kikuno R.-F., Ohara R., Inamoto S., Koseki H., Hirooka S.,  
Saga Y., Seino S., Nishimura M., Katsuo T., Hoshino K., Kitamura H.,  
Nagase T., Ohara O., Koga H.,  
RT "Prediction of the Coding Sequences of Mouse Homologues of KIAA Gene:  
RT IV. The Complete Nucleotide Sequences of 500 Mouse KIAA-Homologous  
RT cDNAs Identified by Screening of Terminal Sequences of cDNA Clones  
RT Randomly Sampled from Size-Fractionated Libraries.",  
RL DNA Res. 11:205-218(2004).  
DR EMBL: AK173112; BAD32390.1;  
DR GO: GO:0005615; C:extracellular space; TAS.  
DR GO: GO:0016021; C:integral to membrane; TAS.  
DR InterPro: IPR001461; Peptidase A1.  
DR InterPro: IPR009120; Pept. A1 BACE1.  
DR InterPro: IPR009007; Pept. Aspartic.  
DR PRINTS: PR01816; BACE1.  
DR PRINTS: PRO07992; PEPSTN.  
FT NON TER 1  
SQ SEQUENCE 373 AA; 41710 MW; 9D7D19D039577732 CRC64;

Query Match 74.8%; Score 1711; DB 2; Length 373;  
Best Local Similarity 98.2%; Pred. No. 2, 3e-129;  
Matches 320; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 108 YVPYTGKKEGELTDLVSIPIHGVTVTVRANIAITSDKPEFINSNEGILGLAVAEIA 167  
DB 1 YVPYTGKKEGELTDLVSIPIHGVTVTVRANIAITSDKPEFINSNEGILGLAVAEIA 60  
QY 168 RPDSDLEPFSDLVKQTHVPLNFSQOLGAGPPLNQSIVLSVGSMTIGIDHSITVGS 227  
DB 61 RPDSDLEPFSDLVKQTHVPLNFSQOLGAGPPLNQSIVLSVGSMTIGIDHSITVGS 120  
QY 228 LMYTIRREMYEVIIIVVEINGQDLKMDCKEYNDKSIIVSGTTLRLPKVFEAAVKS 287  
DB 121 LMYTIRREMYEVIIIVVEINGQDLKMDCKEYNDKSIIVSGTTLRLPKVFEAAVKS 180  
QY 288 IKAASSTKFPDGFGLGEOLVCMOAGTTPMNIPIVLSLIMGEVTVNSGFRITTLIPQOYL 347  
DB 181 IKAASSTKFPDGFGLGEOLVCMOAGTTPMNIPIVLSLIMGEVTVNSGFRITTLIPQOYL 240  
QY 348 FVEDVATISQDCCYKRAISQSSGTGTGAVIMEGFYVFPDRAKRGFAVSAACHVDEPRT 407  
DB 241 FVEDVATISQDCCYKRAISQSSGTGTGAVIMEGFYVFPDRAKRGFAVSAACHVDEPRT 300  
QY 408 AAVEGPFVTLDMEDCGNINIPQDEST 433

DB 301 AAVEGPFVTLDMEDCGNINIPQDEST 326

## RESULT 11

Q9CUC5 PRELIMINARY; PRT; 267 AA.  
AC 09CUC5: 01-JUN-2001 (TREMBlrel. 17, Created)  
DT 01-JUN-2001 (TREMBlrel. 23, Last sequence update)  
DE 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
DE Mus musculus adult male brain cDNA, RIKEN full-length enriched  
DE library, clone:3526402A15 product:beta-site APP cleaving enzyme, full  
DE insert sequence. (Fragment).  
GN Name=Bace1; Synonyms=Bace;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCB1\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Brain;  
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;  
RA Carrini P., Hayashizaki Y.,  
RT "High-efficiency full-length cDNA cloning.",  
RL Meth. Enzymol. 303:19-44(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Brain;  
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;  
RA RIKEN PANTOM Consortium;  
RT "Functional annotation of a full-length mouse cDNA collection.",  
RL Nature 409:685-690(2001).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Brain;  
RA The PANTOM Consortium;  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs.",  
RL Nature 420:563-573(2002).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Brain;  
RX MEDLINE=2049374; PubMed=11042159; DOI=10.1101/gr.145100;  
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
RT prepare full-length cDNA libraries for rapid discovery of new genes.",  
RL Genome Res. 10:1617-1630(2000).  
RN [5]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Brain;  
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,  
RA Komori H., Akiyama J., Nishi K., Kitanishi T., Tashiro H., Itoh M.,  
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ilegami T., Kashiwagi K.,  
RA Fujiwara S., Inoue K., Togawa Y., Ozawa K., Tanaka T., Matsunaga S., Kawai J.,  
RA Yoneda Y., Ishikawa T., Ogasawara Y., Iwata M., Ohara E., Watanabe M.,  
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,  
RT "RIKEN integrated sequence analysis (RISA) system-384-format  
RT sequencing pipeline with 384 multicapillary sequencer.",  
RL Genome Res. 10:1757-1771(2000).  
RN [6]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Brain;  
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,  
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,  
RA Haragaki T., Hara A., Hayatsu N., Hiramoto K., Hirooka T., Hori F.,  
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,  
RA Kawai J., Kojima Y., Komori H., Kouda M., Koya S., Kuzuhara C.,  
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,



RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,  
 RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,  
 RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,  
 RA Tejima Y., Toyota T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,  
 RA Muramatsu M., Hayashizaki Y.;  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK014390; BAB29317.2; -.  
 DR HSSP; P56817; 1FKN.  
 DR MGD; MGI:1346542; Bace1.  
 DR GO; GO:0005615; C:extracellular space; TAS.  
 DR GO; GO:0016021; C:integral to membrane; TAS.  
 DR InterPro; IPR009120; Pept\_A1\_BACE1.  
 DR InterPro; IPR009007; Pept\_Aspartic.  
 DR PRINTS; PRO1816; BACE1.  
 DR FT NON TER 1 1  
 SQ SEQUENCE 267 AA; 3033 MW; 9413BA4530AB63B0 CRC64;  
 Query Match 50.8%; Score 1163; DB 2; Length 267;  
 Best Local Similarity 99.1%; Pred. No. 2, 1e-85;  
 Matches 218; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 214 MIIIGDHSLYTGSMTYTPRRMYEVIIVRVEINGQDLKMDCKEYNDKSIDSGTTN 273  
 DB 1 MIIIGDHSLYTGSMTYTPRRMYEVIIVRVEINGQDLKMDCKEYNDKSIDSGTTN 60  
 QY 274 LRLPKVFPEAFAVKSIAASSTKFPDGFMLGEQVCMQAGTTFWNIIPVILSYLMGEVTN 333  
 DB 61 LRLPKVFPEAFAVKSIAASSTKFPDGFMLGEQVCMQAGTTFWNIIPVILSYLMGEVTN 120  
 QY 334 QSFRRITLPPQYLRPEVDVATSDDCYKRAISGSGTGTWGANVINGFYVPPRARRIG 393  
 DB 121 QSFRRITLPPQYLRPEVDVATSDDCYKRAISGSGTGTWGANVINGFYVPPRARRIG 180  
 QY 394 FAVSACHVDEPRTAIVEGPFVTLDMEDGYNIPQIDEST 433  
 DB 181 FAVSACHVDEPRTAIVEGPFVTLDMEDGYNIPQIDEST 220

RESULT 12  
 BAE2\_HUMAN STANDARD; PRT; 518 AA.  
 ID BAE2\_HUMAN STANDARD; PRT; 518 AA.  
 AC Q95Z0; Q9UT6;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Beta secretase 2 precursor (BC 3.4.23.45) (Beta-site APP-cleaving  
 DE enzyme 2) (Aspartyl protease 1) (Asp 1) (ASPL) (Membrane-associated  
 DE aspartic protease 1) (Memapsin-1) (Down region aspartic protease)  
 DE (UNQ418/PRO852).  
 DE GN Name=BACE2; Synonyms=ASP21;  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20057170; PubMed=10591213; DOI=10.1038/990107;  
 RA Yan R., Bienkowski M.J., Shuck M.E., Miao H., Torg M.C., Pauley A.M.,  
 RA Braisher J.R., Strattan N.C., Mathews W.R., Buhl A.E., Carter D.B.,  
 RA Tomaselli A.G., Parodi L.A., Heinrichson R.L., Gurney M.B.;  
 RT "Membrane-anchored aspartyl protease with Alzheimer's disease beta-  
 RT secretase activity.";  
 RL Nature 402:533-537(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Bone marrow;  
 RA Xin H., Stephens J.C., Duan X., Harrowe G., Kim E., Grieshammer U.,  
 RA Giese K.;  
 RT "Identification of a novel aspartic-like protease differentially  
 RT expressed in human breast cancer cell lines";  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.

RA Accarino M.P., Pumagalli P., Ottolenghi S., Taramelli R.;  
 RT "Cloning of a gene from chromosome 21 Down region encoding a potential  
 RT transmembrane aspartyl protease.";  
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Solans A., Estivill X., de la Luna S.;  
 RT "Cloning of a novel mammalian aspartyl protease.";  
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20120043; PubMed=10656250; DOI=10.1006/mcne.1999.0811;  
 RA Huseain I., Powell D.J., Howlett D.R., Tew D.G., Meek T.D.,  
 RA Chapman C., Gloger I.S., Murphy K.E., Southern C.D., Ryan D.M.,  
 RA Smith T.S., Simmons D.L., Walsh F.S., Dingwall C., Christie G.;  
 RT "Identification of a novel aspartic proteinase (Asp 2) as beta-  
 RT secretase.";  
 RL Mol. Cell. Neurosci. 14:419-427(1999).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2014060; PubMed=10677483; DOI=10.1073/pnas.97.4.1456;  
 RA Lin X., Koelsch G., Wu S., Downs D., Danti A., Tang J.;  
 RT "Human aspartic protease memapsin 2 cleaves the beta-secretase site of  
 RT beta-amyloid precursor protein.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 97:1456-1460(2000).  
 RN [7]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22887396; PubMed=12975309; DOI=10.1101/gr.1293003;  
 RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D., Brush J.,  
 RA Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,  
 RA Eaton D., Foster J., Grimaldi C., Gu Q., Haas P.B., Heldens S.,  
 RA Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,  
 RA Lewis L., Liao D., Mark M., Robble E., Sanchez C., Schoenfeld J.,  
 RA Seshagiri S., Simons L., Singh J., Smith V., Stinson J., Vagts A.,  
 RA Vandlen R., Watanabe C., Wleand D., Woods K., Xie M.-H., Yanura D.,  
 RA Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A., Wood W.I.,  
 RA Godowski P., Gray A.;  
 RT "The secreted protein discovery initiative (SPDI), a large-scale  
 RT effort to identify novel human secreted and transmembrane proteins: a  
 RT bioinformatics assessment.";  
 RL Genome Res. 13:2265-2270(2003).  
 RN [8]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20289799; PubMed=10830953; DOI=10.1038/35012518;  
 RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,  
 RA Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Soeda E.,  
 RA Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,  
 RA Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,  
 RA Reichwald A., Rump J., Shibuya K., Kawasaki K., Asakawa S.,  
 RA Renschthal A., Kudoh J., Shibuya K., Kawasaki K., Antonarakis S.E.,  
 RA Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Brandt P.,  
 RA Minoshima S., Shimizu N., Nordstiek G., Hornischer K., Brandt P.,  
 RA Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloecher H.,  
 RA Ramer J., Beck A., Klages S., Hennig S., Riessmann L., Dagand B.,  
 RA Wehner S., Borzym K., Gardiner K., Nizetic D., Francis F.,  
 RA Leirich H., Reinhardt R., Yaspo M.-L.;  
 RT "The DNA sequence of human chromosome 21.";  
 RL Nature 405:311-319(2000).  
 RN [9]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Skin  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klauenberg R.D., Collins F.S., Wagner F.S., Shenmen C.M., Schuler G.D.,  
 RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Hecht N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hecht F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,  
 RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loughellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McKernan P.J., McKernan K.J., Malek J.G., Gundaretan P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shenchenko Y., Bouffard G.G.,  
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [10]  
RP CHARACTERIZATION.  
RX MEDLINE:22088158; PubMed:12093293; DOI=10.1021/bi025926t;  
RA Turner R.T. III, Loy J.A., Nguyen C., Devaasudram T., Ghosh A.K.,  
RA Koelsch G., Tang J.,  
RT "Specificity of memapsin 1 and its implications on the design of  
RT memapsin 2 (beta-secretase) inhibitor selectivity.";  
RL Biochemistry 41:8742-8746(2002).  
CC -1- CATALYTIC ACTIVITY: Broad endopeptidase specificity. Cleaves Glu-  
CC Val-Aen-Leu-1-Asp-Ala-Glu-Phe in the Swedish variant of  
CC Alzheimer's amyloid precursor protein.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- SIMILARITY: Belongs to the peptidase A1 family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AF200342; AAF17078.1; -  
DR EMBL; AF117882; AAD45240.1; -  
DR EMBL; AF050171; AAD45963.1; -  
DR EMBL; AF178532; AAF29494.1; -  
DR EMBL; AF204944; AAF26368.1; -  
DR EMBL; AF200192; AAF13714.1; -  
DR EMBL; AY358927; AAO89286.1; -  
DR EMBL; AL163284; CAB90458.1; -  
DR EMBL; BC014453; CAB90554.1; -  
DR HSSP; P56817; 1M4H.  
DR MEROPS; A01.041; -  
DR Genew; HGNC:934; BACE2.  
DR MIM; 605668; -  
DR GO; GO:0005624; C:membrane fraction; TAS.  
DR GO; GO:0004190; F:aspartic-type endopeptidase activity; TAS.  
DR GO; GO:0006464; P:protein modification; TAS.  
DR GO; GO:0009306; P:protein secretion; TAS.  
DR InterPro; IPR009119; Pept\_A1\_BACE.  
DR InterPro; IPR009121; Pept\_A1\_BACE2.  
DR InterPro; IPR009007; Pept\_Asp\_AS.  
DR InterPro; IPR001461; Peptidase\_A1.  
DR Pfam; PF00026; Asp\_1.  
DR PRINTS; PR01817; BACE2.  
DR PRINTS; PR01815; BACEFAMILY.  
DR PROSITE; PS00792; PEPsin.  
DR PROSITE; PS00141; ASP\_PROTEASE; 2.  
KW Aspartyl protease; Glycoprotein; Hydrolase; Signal; Transmembrane;  
KW Zymogen.  
FT SIGNAL 1 20 Potential.  
FT PROPEP 21 ? Potential.  
FT CHAIN 21 518 Beta secretase 2.  
FT DOMAIN 21 473 Extracellular (Potential).  
FT TRANSMEM 474 494 Potential.  
FT DOMAIN 495 518 Cytoplasmic (Potential).  
FT ACT\_SITE 110 110 By similarity.  
FT ACT\_SITE 303 303 By similarity.  
FT CARBOHYD 170 170 N-linked (GlcNAc...) (Potential).  
FT CARBOHYD 366 366 N-linked (GlcNAc...) (Potential).  
FT CONFLICT 36 36 A -> T (in Ref. 6).  
SQ SEQUENCE 518 AA; 56180 MW; 2B903150823760D3 CRC64;

Query Match 49.7%; Score 1136; DB 1; Length 518;  
Best Local Similarity 52.2%; Pred. No. 7,8e-83;  
Matches 215; Conservative 66; Mismatches 121; Indels 10; Gaps 3;  
QY 8 PLRSLGAGPLGLINLETDERPEGRGRGSFVEMVDNLRGKSGQGVVENTVSPPTL 67  
DB 52 PARRHADGLALALPALAS-----PAGANPFLMVDLQDSGRGYLEMLIGPPQKL 105  
QY 68 NILVDTGSSNFAVGAAPPEFLHRYIQRLSTYDLRKGVVPTYGKMEGLCTDLVSI 127  
DB 106 QLVDTGSSNFAVGAAPPEFLHRYIQRLSTYDLRKGVVPTYGKMEGLCTDLVSI 165  
QY 128 PRGPVTVRANIAATTSDEKFFNGSNMGLGAAYEIAAPDSELPFDSLVKQTHVP 187  
DB 166 PKGFPTSLVNIATIFESSENFLLGKNGILGAIYTLAPSSSLTFPDSLVTQANIP 225  
QY 188 NLFSLQCGAGPPLNQSEVLASVQGSMTIGGIDHSILYTSGLWYTPIRREYEVYIVRVE 247  
DB 226 NVFSMQMGAGIPVAGS---GTNGSILVGGIEPSLYKGDIWYTPKEEYVYQIEILKLE 282  
QY 248 INGQDLKMDCKEYVYVDSGTTLRLPKVFEAAVKSIKAASTKRPDGFMLGEOL 307  
DB 263 IGGQSLNIDCKEYVYVDSGTTLRLPKVFEAAVKSIKAASTKRPDGFMLGEOL 342  
QY 308 VQWQAGTTPMNIPEVISLYMGEVYVDSGTTLRLPKVFEAAVKSIKAASTKRPDGFMLGEOL 367  
DB 343 ACWTNSETPWSYFPKISILYRDENSRSFRITLLPOLYIQPMNAGLANI-ECRFRFISFS 401  
QY 368 STGTVMGAVMEGVYVFDRAKRIKGAIVSACHVBEFRTAAVEGPFVTLDM 419  
DB 402 TNAIVTATWEGHYVFDRAKRIKGAIVSACHVBEFRTAAVEGPFVTLDM 453  
RESULT 13  
ID Q9H2V8 PRELIMINARY; PRT; 439 AA.  
AC Q9H2V8;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE CDA13.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Pnechocytoma;  
RA Li Y., Huang Q., Peng Y., Song H., Yu Y., Xu S., Ren S., Chen Z.,  
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: Belongs to peptidase family A1.  
DR EMBL; AF212252; AAG41783.1; -  
DR HSSP; P56817; 1FXN.  
DR GO; GO:0016021; C:integral to membrane; ISS.  
DR GO; GO:0006509; P:membrane protein ectodomain proteolysis; ISS.  
DR GO; GO:0042985; P:negative regulation of amyloid precursor pr...; ISS.  
DR GO; GO:0016486; P:peptide hormone processing; ISS.  
DR InterPro; IPR001461; Peptidase\_A1.  
DR InterPro; IPR009119; Pept\_A1\_BACE.  
DR InterPro; IPR009121; Pept\_A1\_BACE2.  
DR InterPro; IPR009007; Pept\_Asp\_AS.  
DR PRINTS; PR01817; BACE2.  
DR PRINTS; PR00792; PEPsin.  
DR PRINTS; PR01815; BACEFAMILY.  
DR PROSITE; PS00141; ASP\_PROTEASE; 2.  
KW Aspartyl protease; Hydrolase; Protease.  
SQ SEQUENCE 439 AA; 48275 MW; 02EC0E050F11602 CRC64;  
Query Match 49.0%; Score 1121; DB 2; Length 439;  
Best Local Similarity 54.8%; Pred. No. 1e-81;

Matches 207; Conservative 63; Mismatches 104; Indels 4; Gaps 2;

QY 42 MVDNLKSGGQGYVEMTVGSPQTLNIVDTGSSNFVGAAPHPLHRYGRLSTYR 101  
 DB 1 MVDNLQGSBGRVYLEMLGTPEPQKQILVDTGSSNFVGAAPHGYIDTYFTERSSTYR 60

QY 102 DIRKGVYVYPTQCKEGEGLTDLVSPHPNPVTVRANIAITESDKPFINGSNWEIITL 161  
 DB 61 SKGPDVTVKYGTSWVGFBEDLVITPKGPNISFLVNITIFESSENFLLPGIKMGITGL 120

QY 169 AYAARIAPDLSLEPFDLSVKQTHVNLNLSLOLCAAGFLPNQSEVLAASVGSMTIGDIDH 221  
 DB 121 AYATIAKSSSLETFFDLSVTQANTINVSMMQCGALFVAGS---GTNGSIVLGIPT 177

QY 222 SIYTGSLWYPIRREWYEVIIIVRVINGQDLKMDCKEYNYKSIYDSGTTNLRPKYF 281  
 DB 178 SIYKGDIVTPIKEBMYQIEILKEITGGQSLNLDREYNADKAIYDSGTTNLRPKYF 237

QY 282 EAAVNSIKRAASTERKPPDGFVIGELVQVQAGTTPNITPVISLYMGVTVMSRITTL 341  
 DB 238 DAVEVAVARASLIPERPSDFGFTGSQLACWTNSETPWSYFPKISIVLRDNSSRSPRITTL 297

QY 342 PQOYLKRVVDVATSGDDCYKFAISOSTGTWGAIVMEGFYVFDRAKRIGPVAACHV 401  
 DB 298 POLYIQQPMWAGALNY-ECTRFGISPSITMALVIGATVMEGFYVFDRAKRVGFAPASPCAE 356

QY 402 HDEFRTAAGEPFTVLTDM 419  
 DB 357 INGAANSEISGPSTEDV 374

RESULT 14  
 Q8CSE9 PRELIMINARY; PRT; 514 AA.

AC Q8CSE9; 01-MAR-2003 (T-EMBLrel. 23, Created)  
 DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)  
 DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)  
 DE muscle 15 days embryo male testis cDNA, RIKEN full-length  
 DE enriched library, clone:8030470009 product:beta-site APP-cleaving  
 DE enzyme 2, full insert sequence.  
 GN Name:Bace2;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Testis;  
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;  
 RA Carninci P., Hayashizaki Y.;  
 RT "High-efficiency full-length cDNA cloning.";  
 RL Meth. Enzymol. 303:19-44(1999).  
 RN [2]  
 RP SQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Testis;  
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;  
 RA RIKEN FANTOM Consortium;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 RN [3]  
 RP SQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Testis;  
 RA The FANTOM Consortium.  
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";  
 RL Nature 420:563-573(2002).  
 RN [4]  
 RP SQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Testis;  
 RX MEDLINE=2049374; PubMed=11042159; DOI=10.1101/gr.145100;  
 RA Carninci P., Shibata Y., Hayatsu N., Suganara Y., Shibata K., Itoh M.,

RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";  
 RL Genome Res. 10:1617-1630(2000).  
 RN [5]  
 RP SQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Testis;  
 RX MEDLINE=20530913; PubMed=11076661; DOI=10.1101/gr.152600;  
 RA Shibata K., Itoh M., Atzawa K., Nagaoka S., Sasaki N., Carninci P.,  
 RA Konno H., Akiyama J., Nishi K., Kitsuana T., Tashiro H., Itoh M.,  
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
 RA Fujiwara S., Inoue K., Togawa M., Izawa M., Ohara E., Watabiki M.,  
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,  
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
 RT "RIKEN integrated sequence analysis (RISA) system-384-format  
 RT sequencing pipeline with 384 multichipillary sequencer.";  
 RL Genome Res. 10:1757-1771(2000).  
 RN [6]  
 RP SQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Testis;  
 RA Adachi J., Atzawa K., Akimura T., Arakawa T., Bono H., Carninci P.,  
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,  
 RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoaka T., Hirozane T.,  
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,  
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,  
 RA Kurihara C., Matsuyama T., Miyazaki A., Mura M., Nakamura M.,  
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,  
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,  
 RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,  
 RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,  
 RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;  
 RL Submitted (Apr-2002) to the EMBL/Genbank/DBJ databases.  
 CC -1- SIMILARITY: Belongs to peptidase family A1.  
 DR EMBL: AK078770; BAC37384.1; -  
 DR HSSP: P56817; 1PKN.  
 DR MEROPS: A01.041; -  
 DR MGD: MG1:160440; Bace2.  
 DR GO: GO:0005615; C:extracellular space; TMS.  
 DR GO: GO:0016021; C:integral to membrane; ISS.  
 DR GO: GO:0006509; P:membrane protein ectodomain proteolysis; ISS.  
 DR GO: GO:0042985; P:negative regulation of amyloid precursor pr. .; ISS.  
 DR GO: GO:0016486; P:peptide hormone processing; ISS.  
 DR InterPro: IPR001461; Peptidase A1.  
 DR InterPro: IPR005119; Pept\_A1-BACE.  
 DR InterPro: IPR009121; Pept\_A1-BACE2.  
 DR InterPro: IPR009007; Pept\_Aspartic.  
 DR InterPro: IPR001969; Pept\_Asp\_AS.  
 DR PRINTS: PRO1817; BACE2.  
 DR PRINTS: PRO1815; BACSFAMILY.  
 DR PRINTS: PRO0792; PEPsin.  
 DR PROSITE: PS00141; ASP\_PROTEASE; 2.  
 DR Aspartyl protease; Hydrolase; Protease.  
 KW SQUENCE 514 AA; 55810 MW; CB99237B86A0B2E CRC64;  
 SO

Query Match 49.0%; Score 1120; DB 2; Length 514;  
 Best Local Similarity 52.5%; Pred. No. 1.5e-81;  
 Matches 217; Conservative 67; Mismatches 113; Indels 16; Gaps 6;

QY 12 GIG--GAP--GLEINLTDEPBERGR-RGSFVENVDLRGKSGGYVEMTVGSPPT 66  
 DB 48 GATGELPRADGALAL-----EPVATNFAMVNLQDSRGVYLEMLIGTPPK 100

QY 67 LNLIVDTGSSNFVGAAPHPLHRYGRLSTYDLRGVYVYPTQCKEGEGLTDLVS 126  
 DB 101 VOIVDTGSSNFVGAAPHPLHRYGRLSTYDLRGVYVYPTQCKEGEGLTDLVS 160

QY 127 IPRGVNTVVRANIAITESDKPFINGSNWEIITGLAIAIAPDLSLEPFSIVQTHV 186  
 DB 161 IPKGFNSFLVNIATIFESSENFLLPGIKWNGILGLAIAIAPDLSLEPFSIVQTHV 220

QY 187 PNLFSIOLGAGPLPNQSEVLAASVGSMTIGGIDSLYGSIMWYPIRREWYEVIIIV 246  
 DB 187 PNLFSIOLGAGPLPNQSEVLAASVGSMTIGGIDSLYGSIMWYPIRREWYEVIIIV 246

Db 221 PDIFSMQCGAGLPVAGS---GTNGSLVLGGIEPSLYKGDWYTPIKEEYVQIETILKL 277

QY 247 EINGODLKMDCKEYNYDKSIYDSGTTNLRPKKVEAAVKSIKAASSTKEPDPGFMLGEO 306

Db 278 EIGGQNLMDCKEYNADRAIVDSGTTLLRLPKQVDAVAEVAARTSLIPESDSGFMTGAQ 337

QY 307 LVCWQAGTTPNNIIPFVLSLYLMEGVNOSFRITLLPOOYLRPVEDVATSDODCYKFAISQ 366

Db 338 LACWTNSETPMAYFPFKISYILRDENASRSFRITLLPOLYIQPMGAGFNY-ECYRFGISS 396

QY 367 SSTGTWGAIVMEGFYVFDRAKRIKIGFPAVSACHVHDEFRTAAVEGPFVTLDM 419

Db 397 STNALVIGATVMEGFYVFDRAQRVGFPAVSPCAIEGTTSISGPFSTEDI 449

RESULT 15

OSJL18 PRELIMINARY; PRT; 514 AA.

ID 09JL18

AC 09JL18;

DT 01-OCT-2000 (Tremblrel. 15, Created)

DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)

DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)

DE Aspartyl protease 1.

GN Name-Bace2;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RA Choi D.K., Sugano S., Sakaki Y.;

RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: Belongs to peptidase family A1.

DR EMBL; AF216310; AAF36599.1; -

DR HSBP; P56817; 1PKN.

DR MEROPS; A01.041; -

DR MGD; MGI:1860440; Bace2.

DR GO; GO:0005615; C:extracellular space; TAS.

DR GO; GO:0016021; C:integral to membrane; ISS.

DR GO; GO:0006509; P:membrane protein ectodomain proteolysis; ISS.

DR GO; GO:0042985; P:negative regulation of amyloid precursor pr. . .; ISS.

DR GO; GO:0016486; P:peptide hormone processing; ISS.

DR InterPro; IPR001461; Peptide\_hormone\_A1.

DR InterPro; IPR009119; Pept\_A1\_BACE.

DR InterPro; IPR009121; Pept\_A1\_BACE2.

DR InterPro; IPR009007; Pept\_Aspartic.

DR InterPro; IPR001969; Pept\_Asp\_AS.

DR PRINTS; PRO1817; BACE2.

DR PRINTS; PRO1815; BACEFAM1LY.

DR PRINTS; PRO0792; PEPSIN.

DR PROSITE; PS00141; ASP\_PROTEASE; 2.

DR Aspartyl protease; Hydrolyase; Protease.

KW SEQUENCE 514 AA; 55799 MW; A70725F2C1DF5B47 CRC64;

Query Match 49.0%; Score 1120; DB 2; Length 514;

Best Local Similarity 52.3%; Pred. No. 1.5e-81; Indels 16; Gaps 5;

Matches 216; Conservative 67; Mismatches 114;

QY 12 GIGGAPL-----GLEINLETDEEPEEGR-RGSFVEMVDNLKSGSGGYVEMTVGSPPO 66

Db 48 GIGTELPRAADIGALAL-----EPVATANFLAMVDNLQDSGSGYVLEMLIGTTPQK 100

QY 67 LNIIVDTGSSNFVAGAPHPFLHRYRQKSLSTYDLKRGYVVPYPTQGMKGEIGTDLVS 126

Db 101 VOIIVDTGSSNFVAGAPHSIYDTYFDESSTYHSKGFDTYVYKTSQSWTGFVGEDLV 160

QY 127 IPHGNVTVRANIAITESDKFINGSNWEGILGLAYAEIARPPDSLEFPFDSLVKQTHV 186

Db 161 IPKGNSSFLVNIATIFESNFPGLGIMNGILGLAYALAKPSSSLETFPDSLVAQAKI 220

QY 187 PNLPSLOLCGAGFPALNGEVLASVGSNIIIGIHSLYTGSLWYTPTRREYVEVILIVR 246

Db 221 PDIFSMQCGAGLPVAGS---GTNGSLVLGGIEPSLYKGDWYTPIKEEYVQIETILKL 277

QY 247 EINGODLKMDCKEYNYDKSIYDSGTTNLRPKKVEAAVKSIKAASSTKEPDPGFMLGEO 306

Db 278 EIGGQNLMDCKEYNADRAIVDSGTTLLRLPKQVDAVAEVAARTSLIPESDSGFMTGAQ 337

QY 307 LVCWQAGTTPNNIIPFVLSLYLMEGVNOSFRITLLPOOYLRPVEDVATSDODCYKFAISQ 366

Db 338 LACWTNSETPMAYFPFKISYILRDENASRSFRITLLPOLYIQPMGAGFNY-ECYRFGISS 396

QY 367 SSTGTWGAIVMEGFYVFDRAKRIKIGFPAVSACHVHDEFRTAAVEGPFVTLDM 419

Db 397 STNALVIGATVMEGFYVFDRAQRVGFPAVSPCAIEGTTSISGPFSTEDI 449

Search completed: August 22, 2005, 00:24:01

Job time : 136 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 22, 2005, 00:26:34 ; Search time 302 Seconds  
(without alignment)  
2346.052 Million cell updates/sec

Title: US-10-726-967a-84

Perfect score: 2288

Sequence: 1 TGGGIRLPLRSLGGAPLGL.....FVTLMEGCGINIPQTDST 433

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xld  
-Q/cgn2\_1/usfto.spool\_p/us10726967/runat.19082005.201241.29654/app.query.fasta\_1.583  
-DB=Issued\_Patents\_NA -OPMT=fastap -SUPRT=rm1 -MINMATCH=0.1 -LOOPT=0  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTPMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10726967 @CGN 1.1 105 @runat.19082005.201241.29654 -NCPU=6 -ICPU=3  
-NO MAP -LARGEOUTERY -NEG\_SCORES=0 -WAIT -DSBLOC=100 -LONGLOG  
-DEV\_TIMEOUT=120 -MAPN\_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Issued Patents NA:\*

- 1: /cgn2\_6/prodata/1/ina/5A.COMB.seq:\*
- 2: /cgn2\_6/prodata/1/ina/5B.COMB.seq:\*
- 3: /cgn2\_6/prodata/1/ina/6A.COMB.seq:\*
- 4: /cgn2\_6/prodata/1/ina/6B.COMB.seq:\*
- 5: /cgn2\_6/prodata/1/ina/PCTUS.COMB.seq:\*
- 6: /cgn2\_6/prodata/1/ina/backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2267	99.1	1503	4	US-09-724-566A-1
2	2267	99.1	1503	4	US-09-471-669A-1
3	2267	99.1	2070	3	US-09-548-372D-3
4	2267	99.1	2070	3	US-09-548-372D-3
5	2267	99.1	2070	3	US-09-548-372D-3
6	2267	99.1	2070	4	US-09-551-853D-3
7	2267	99.1	2070	4	US-09-416-901B-3
8	2267	99.1	2070	4	US-09-548-372D-3
9	2267	99.1	2070	4	US-09-794-927A-3
10	2267	99.1	2070	4	US-09-548-372D-3
11	2267	99.1	2070	4	US-09-795-847B-3
12	2267	99.1	2070	4	US-09-548-366F-3

13	2267	99.1	2070	4	US-09-794-925A-3	Sequence 3, Appl1
14	2267	99.1	2370	4	US-09-806-194A-3	Sequence 3, Appl1
15	2267	99.1	2048	4	US-09-724-566A-42	Sequence 42, Appl1
16	2267	99.1	2348	4	US-09-724-566A-44	Sequence 44, Appl1
17	2267	99.1	2348	4	US-09-471-669A-42	Sequence 42, Appl1
18	2267	99.1	2348	4	US-09-471-669A-44	Sequence 44, Appl1
19	2267	99.1	3252	4	US-09-604-608-1	Sequence 1, Appl1
20	2267	99.1	5825	4	US-09-949-016-4382	Sequence 4382, Ap
21	2267	99.1	5878	4	US-09-949-016-939	Sequence 939, Ap
22	2267	99.1	16080	4	US-09-724-566A-48	Sequence 48, Appl1
23	2267	99.1	16080	4	US-09-471-669A-48	Sequence 48, Appl1
24	2267	99.1	2070	4	US-09-869-414-3	Sequence 3, Appl1
25	2267	99.1	2541	3	US-09-009-191-1	Sequence 1, Appl1
26	2257	98.6	1302	3	US-09-548-372D-25	Sequence 25, Appl1
27	2257	98.6	1302	3	US-09-548-372D-25	Sequence 25, Appl1
28	2257	98.6	1302	3	US-09-551-853D-25	Sequence 25, Appl1
29	2257	98.6	1302	4	US-09-416-901B-25	Sequence 25, Appl1
30	2257	98.6	1302	4	US-09-548-376D-25	Sequence 25, Appl1
31	2257	98.6	1302	4	US-09-794-927A-25	Sequence 25, Appl1
32	2257	98.6	1302	4	US-09-548-373D-25	Sequence 25, Appl1
33	2257	98.6	1302	4	US-09-795-847B-25	Sequence 25, Appl1
34	2257	98.6	1302	4	US-09-869-414-25	Sequence 25, Appl1
35	2257	98.6	1302	4	US-09-548-366F-25	Sequence 25, Appl1
36	2257	98.6	1302	4	US-09-548-368D-25	Sequence 25, Appl1
37	2257	98.6	1302	4	US-09-794-925A-25	Sequence 25, Appl1
38	2257	98.6	1302	4	US-09-806-194A-25	Sequence 25, Appl1
39	2257	98.6	1341	3	US-09-548-372D-21	Sequence 21, Appl1
40	2257	98.6	1341	3	US-09-548-367D-21	Sequence 21, Appl1
41	2257	98.6	1341	4	US-09-551-853D-21	Sequence 21, Appl1
42	2257	98.6	1341	4	US-09-416-901B-21	Sequence 21, Appl1
43	2257	98.6	1341	4	US-09-548-376D-21	Sequence 21, Appl1
44	2257	98.6	1341	4	US-09-794-927A-21	Sequence 21, Appl1
45	2257	98.6	1341	4	US-09-548-373D-21	Sequence 21, Appl1

#### ALIGNMENTS

RESULT 1  
US-09-724-566A-1  
Sequence 1, Application US/09724566A  
Patent No. 6627739  
GENERAL INFORMATION:  
APPLICANT: Anderson, John P.  
APPLICANT: Basl, Gurigbal  
APPLICANT: Doane, Minh Tam  
APPLICANT: Frigon, No. 6627739mand  
APPLICANT: John, Varghese  
APPLICANT: Power, Michael  
APPLICANT: Srinha, Sukanto  
APPLICANT: Tateono, Gwen  
APPLICANT: Tung, Jay  
APPLICANT: Wang, Shuwen  
TITLE OF INVENTION: Beta-Secretase Enzyme Compositions and  
TITLE OF INVENTION: Methods  
FILE REFERENCE: 228-US-NWC2  
CURRENT APPLICATION NUMBER: US/09/724,566A  
CURRENT FILING DATE: 2000-11-28  
PRIOR APPLICATION NUMBER: US 09/501,708  
PRIOR FILING DATE: 2000-02-10  
PRIOR APPLICATION NUMBER: 60/119,571  
PRIOR FILING DATE: 1999-02-10  
PRIOR APPLICATION NUMBER: 60/139,172  
PRIOR FILING DATE: 1999-06-15  
NUMBER OF SEQ ID NOS: 104  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1  
LENGTH: 1503  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-724-566A-1

Alignment Scores:

Pred. No.: 2,59e-257 Length: 1503  
 Score: 2267.00 Matches: 429  
 Percent Similarity: 99.31% Conservative: 1  
 Best Local Similarity: 99.08% Mismatches: 3  
 Query Match: 99.08% Indels: 0  
 DB: 4 Gaps: 0

US-10-726-967A-84 (1-433) x US-09-724-566A-1 (1-1503)

QY 1 ThrGlnHisGlyIleArgLeuProLeuArgSerGlyLeuGlyValAlaProLeuGlyLeu 20  
 DB 64 ACCGACGACGGCATCGGCTGCGCGACGGGCTGGGGGGGCGCCCTGGGGCTG 123  
 QY 21 GluIleAsnLeuGlyThrAspGluGlyProGluGlyProGlyArgGlySerPheVal 40  
 DB 124 CGGCTGCGCGGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 183  
 QY 41 GluMetValAspAsnLeuArgGlyLeuSerGlyGlnGlyValAlaGluMetThrVal 60  
 DB 184 GAGATGGTGGACACCTGAGGGGCAAGTCGGGGCAGGGCTTACGTGGAGATGACCGTG 243  
 QY 61 GlySerProGlnThrLeuAsnIleLeuValAspThrGlySerSerAsnPheAlaVal 80  
 DB 244 GCGACGCGCGCGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 303  
 QY 81 GlyAlaAlaProHisProPheLeuHisArgTyrTyrGlnArgGlnLeuSerSerThrTyr 100  
 DB 304 GGTGCTGCG 363  
 QY 101 ArgAspLeuArgGlyValAlaTyrValProTyrThrGlnGlyValTyrGluGlyLeu 120  
 DB 364 CGGACCTCGCGGAGGGGTGTATGTCCTTACACCGACGCGCAAGTGGGAGGGAGCTG 423  
 QY 121 GlyThrAspLeuValSerIleProHisGlyProAsnValThrValArgAlaAsnIleAla 140  
 DB 424 GGCACCGACCTGGTACATCCCCACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 483  
 QY 141 AlaIleThrGlnSerArgPhePhePheIleAsnGlySerAsnTyrGluGlyLeuGly 160  
 DB 484 GCGATCATGTAATCAGACAGATTTCTTATCAGCGCTCCAACTGGGAGGAGCTCTGGGG 543  
 QY 161 LeuAlaTyrAlaGluIleAlaTyrProAspAspSerLeuGlyProPheAspSerLeu 180  
 DB 544 CTGGCTGATGCTAGATTCGACGGCTGACGACCTCTTGGAGCTTTCTTATGCTCTGG 603  
 QY 181 ValIleGlnThrHisValProAsnLeuPheSerLeuGlnLeuGlyAlaGlyPhePro 200  
 DB 604 GTAAACGACACCCACCTTCCAACTCTCTGCGACCTTGTGGTGGCTGGCTTCCC 663  
 QY 201 LeuAsnGlnSerGlyValLeuAlaSerValGlySerMetIleGlyGlyIleAsp 220  
 DB 664 CTCAACCATCTGAAGTGTGCTGCTGCTGCGAGGAGATGATGATGGAGATTCAC 723  
 QY 221 HisSerLeuTyrThrGlySerLeuTyrTyrThrProIleArgArgGluTyrTyrGlu 240  
 DB 724 CACTGCTGTACACAGGAGCTCTGTGTATACACCATCCGCGGAGATGATATATGAG 783  
 QY 241 ValIleIleValAlaGluValGluIleAsnGlyGlnAspLeuGlyMetAspCysLeuGly 260  
 DB 784 GTGATCATTTGGGGGTGGAGATCAATGACAGATCTGAAATGAGCTGCAGAGATAC 843  
 QY 261 AsnTyrAspGlySerIleValAspSerGlyThrThrAsnLeuArgLeuProGlyVal 280  
 DB 844 AACCTATGACAGAGCATTTGTGGACGTGGACACCAACCTTCTTGGCCCAAGAAAGTG 903  
 QY 281 PheGluAlaAlaValIleSerIleValAlaAspSerThrGluValPheProAspGly 300  
 DB 904 TTGAAAGTGGACGTCAAATCATCAAGGACGCTCTTCAAGGAGAGTTCCCTGATGCT 963  
 QY 301 PheTyrLeuGlyGluGlnLeuValCysTyrGlnIleGlyThrThrProTyrPheIlePhe 320  
 DB 964 TTCTGGCTAGAGAGGACGCTGTGTCTGCGACAGGACGACCACTCTTGGAACTTTTTC 1023

QY 321 ProValIleSerLeuTyrLeuMetGlyGluValAlaThrAsnGlnSerPheAspIleThrIle 340  
 DB 1024 CCAAGTATCTTCACTTCACTTAATGGGTGAGTTACCAACGATCTTCCGATCACCATC 1083  
 QY 341 LeuProGlnGlnTyrLeuArgProValGluAspValAlaThrSerGlnAspAspCysTyr 360  
 DB 1084 CTTCGCGACATATCACTCGGCGCAGTGGAGATGGGCACTGCCAAGACGACTTAC 1143  
 QY 361 LysPheAlaIleSerGlnSerSerThrGlyThrValMetGlyAlaValIleMetGlyGly 380  
 DB 1144 AAGTTTCCATCTTCACTCAATCCACGCGCATGTTATGGAGCTTTATCATGAGGGGC 1203  
 QY 381 PheTyrValAlaPheAspArgAlaArgIleArgIleGlyPheAlaValSerAlaCysHis 400  
 DB 1204 TTCTACGTTGCTTGTGATCGGGCGGAAAGCAATTTGGCTGTGTCACGCGCTTGCAT 1263  
 QY 401 ValHisAspGluPheAspArgThrAlaAlaValGluGlyProPheValThrLeuAspMetGly 420  
 DB 1264 GTGCACGATGAGTTCAGACCGCGCGGTGGAGGCGCTTTGTGCACCTTGACATGGAA 1323  
 QY 421 AspCysGlyTyrAsnIleProGlnThrAspGluSerThr 433  
 DB 1324 GACTGTGCTTACATTCACATTCACAGACGATGATCAACC 1362

RESULT 2  
 US-09-471-669A-1  
 / Sequence 1, Application US/09471669A  
 / Patent No. 6830918  
 / GENERAL INFORMATION:  
 / APPLICANT: Anderson, John P.  
 / APPLICANT: Basi, Guribael  
 / APPLICANT: Doane, Minh Tam  
 / APPLICANT: Frigon, No. 6830918mand  
 / APPLICANT: John, Varghese  
 / APPLICANT: Power, Michael  
 / APPLICANT: Sinha, Sukanto  
 / APPLICANT: Tatsuno, Gwen  
 / APPLICANT: Tung, Jay  
 / APPLICANT: Wang, Shuwen  
 / APPLICANT: McConlogue, Lisa  
 / APPLICANT: Eian Pharmaceuticals, Inc.  
 / TITLE OR INVENTION: BETA-SECRETASE ENZYME COMPOSITIONS AND METHODS  
 / FILE REFERENCE: 015270-006430US  
 / CURRENT APPLICATION NUMBER: US/09/471, 669A  
 / CURRENT FILING DATE: 1999-12-24  
 / PRIOR APPLICATION NUMBER: US 60/114, 408  
 / PRIOR FILING DATE: 1998-12-31  
 / PRIOR APPLICATION NUMBER: US 60/119, 571  
 / PRIOR FILING DATE: 1999-02-10  
 / PRIOR APPLICATION NUMBER: US 60/139, 172  
 / NUMBER OF SEQ ID NOS: 108  
 / SOFTWARE: PatentIn Ver. 2.1  
 / SEQ ID NO 1  
 / LENGTH: 1503  
 / TYPE: DNA  
 / ORGANISM: Homo sapiens  
 / US-09-471-669A-1

Alignment Scores:  
 Pred. No.: 2,59e-257 Length: 1503  
 Score: 2267.00 Matches: 429  
 Percent Similarity: 99.31% Conservative: 1  
 Best Local Similarity: 99.08% Mismatches: 3  
 Query Match: 99.08% Indels: 0  
 DB: 4 Gaps: 0

US-10-726-967A-84 (1-433) x US-09-471-669A-1 (1-1503)

QY 1 ThrGlnHisGlyIleArgLeuProLeuArgSerGlyLeuGlyValAlaProLeuGlyLeu 20  
 DB 64 ACCGACGACGGCATCGGCTGCGCGACGGGCTGGGGGGGCGCCCTGGGGCTG 123

QY	21	GIuIIeaenLeuGIuThrAspGIuGIuProGIuGIuProGIaYrGaYgLYsePheVal	40
DB	124	CGCGTCCCCCGGAGACCGACGAAGAGCCGACGAGGCCGCGGAGGGGCAgCTTGTGTG	183
QY	41	GIuWeTValaApnaenLeuArgLYlyseReGlyyGInGlyTYTYrValaGIuWeTThVal	60
DB	184	GAGATGGTGGCAACCTGAGGGGCGCAAGTGGGGCGAGGCTACTACGTGGGAATACCGGTG	243
QY	61	GIYsePProPGInThrLeuasnIleuValAspThrGIYseSerAsnPhealVal	80
DB	244	GCGACCCCCCGGACAGCCCTCAACACTCTGGTGATACAGGCAACAGTAATTGGCAGTG	303
QY	81	GIYlaaIaIaProHIsPProPheLeuHIsArgTYTYrGIaYgInLeuSeSerThTYr	100
DB	304	GGTGCTGCCCCCAACCCCTCTCGCATCGTACTACAGAGGCAAGCTGTCCAGACATATAC	363
QY	101	ArgApLeuArgLYyagLYValTYrValaProTYrThrInGInGlyyETTPGInGlyGIuLeu	120
DB	364	CGGAGCCTCCGAGAGGGTGTATGTGGCTTACACCCAGGGCAAGTGGAGAGGGGAGGTG	423
QY	121	GIYThrApLeuValSerIleProHIsGIYProAsnValThrValaArgIaAsnIleVal	140
DB	424	GCGACCGACTGTGTAGATCCCCCATGGCCCCCAACGTCACTGTGCGGCAACATTGCT	483
QY	141	AlaIleThrGIuSeAspLYaPheHelleaenGIYseSerAsnTPPGInGlyIleuGly	160
DB	484	GCGATCATCGAATCAGACAAgTTCTTACTCAACGGCTCCAACTGGGAAGCATCTGGGG	543
QY	161	LeuAlaTYrAlaGIuIleAlaArgProAspAspSerLeuGIuProPheAspSerIleu	180
DB	544	CTGGCGCTTGTGGAATGGCCAGGGCTTACAGCTCTCCGAGGCTTTCTTTGACTCTGTG	603
QY	181	ValLYsGIuThHIsValaProAsnLeuPheSerLeuGIuLeuCyGlyValaGIYpHePro	200
DB	604	GTHAAGCAAGCCACGTTCCCACTCTTCTCCCTGCAgCTTGTGTGGTGGCTTCCC	663
QY	201	LeuAsnGIuSeGIuValleuIaSeValGIYGIYseMetIleIeGIYGIYIleAsp	220
DB	664	CTCAACAGTCTGAAGTGTGGCTCTGTGGAGGAGCAAGATCATGTAGGATCCAC	723
QY	221	HIsSerLeuTYrThrGIYseSerLeuTPPYrThrProIleArgArgGIuTPPYrTYrGIu	240
DB	724	CAGTGGCTGTACAGGCAgTCTGTGTATACCACTCCGGGGAGATGGATATATAG	783
QY	241	ValIleIleValaArgValaGIuIleAsnGIYInleAsnLeuLYsMeAspCyIeagLYuTYr	260
DB	784	GTHATCATTTGTGGGTGAGATCAATGACAGATCTGAATATGATCGCAAGAGATAC	843
QY	261	AsnTYrAspLYseSerIleValaAspSerGIYThrThraenLeuArgLeuProLYsVala	280
DB	844	AACTATGACAAAGCATTTGTGACAGTGGCAACCAACACTTCGTTCCCAAGAAATGTG	903
QY	281	PheGIuAlaIaIaValLYsSerIleLYsAlaIaAserSerThrGIuLYsPheProAspLY	300
DB	904	TTTGAAGTCGACGTCAATCAATCAAGGAGGCTCTTCCACGGAAGATTCCTGTATGCT	963
QY	301	PheTPPLeuGIYGIuGIuLeuValaCyETTPGIaIaGIYThrThrProTPHAsnIlePhe	320
DB	964	TTCTGGCTTGGAGAGAGCACTGTGTGTCTGGCAAGCAAGCAACCCCTTGAACATTTTC	1023
QY	321	ProValIleSerLeuTYrLeuWeTGIYValaThraenGIuSePheArgIleThrIle	340
DB	1024	CCAGTCATCTCACTTAAATGGGTGAGTTTACCAACAGTCTCTCGCATTCACATTC	1083
QY	341	LeuProGIuGIuTYrIleuArgProValaGIuAspValaIaThSeSerGIaAspAspTYr	360
DB	1084	CTTCCGACAAATACCTGGCGGCAAGGAAAGATGGCCACATCCCAAGACATGTTATC	1143
QY	361	LYsPheAlaIleSerGIuSeSerThrGIYThrValaMetGIYAlaValaIleWeTGIuGIY	380
DB	1144	AAgTTTGCCATCTCAACAGTCAATCCACGGGCACTGTTATGGAGAGCTGTTATCATGAGAGGC	1203
QY	381	PheTYrValaValaPheAspArgIaArgLYsArgIleGIYPheAlaValaSerAlaCyHis	400

[illegible]

Db 424 GGACGACCTGTTAGCATCCCATGCCCCCAACGATCTGTCGGCCAAACATTCCT 483  
Qy 141 AAlIeThrGluSerAspLysPhePheIleAengIySerAsnTrpGluGlyIleuGly 160  
Db 484 GCCATCACTGAATCAAGACAAAGTTCTTCATCAACGGCTCCAACTGGGAAGGATCCTGGG 543  
Qy 161 LeuAlaTrpAlaGluIleAlaArgProAspAspSerIleuGluProPhePheAspSerIleu 180  
Db 544 CTGGCCTATGCTGAATTCAGGAGCTTGAACGATCCGCTCGAGAGCTTCCTTGTGACCTCTG 603  
Qy 181 ValIyGlnThrHisValProAsnLeuPheSerLeuGlnLeuCySGlyAlaGlyPhePro 200  
Db 604 GTAAAGCAGACCCAGTCCCAACCTCTTCCTCCGAGCCTTGTGGTGGCTGGCTTCCC 663  
Qy 201 LeuAengIlnSerGluValIleuAlaSerValGlyIySerMetIleIleGlyIyIleAsp 220  
Db 664 CTCACACAGCTGTAAGTCTGGCTCTGTGAGAGGAGCATGATCATTTGGAGGTATCGAC 723  
Qy 221 HisSerLeuTrpTrpGlySerLeuTrpTrpTrpProIleArgArgGluTrpTrpTrpGlu 240  
Db 724 CACTCGCTGTACACAGGAGCTCTGTGATACCCATCCGCGGAGTGTGATATAG 783  
Qy 241 ValIleIleValArgValGluIleAengIyGlnAspLeuIySerAspCySlyGlyIy 260  
Db 784 GTCACTATTGCGGGTGGAGATCATGACAGGATCTGAAATGCACTCGAAGGAGTAC 843  
Qy 261 AsnTrpAspLysSerIleValAspSerGlyTrpTrpAsnLeuArgLeuProIySlyVal 280  
Db 844 AACTATGACAGACATTTGTGACAGTGGACCCACCACTTGTTGGCCAAAGAAAGT 903  
Qy 281 PheGluAlaAlaValIySerIleValAlaSerSerTrpGluIySlyPheProAspGly 300  
Db 904 TTTAACCTGAGCAATCCATCCAGGAGCTCCCTCCAGGAGAAATTCCTGATGT 963  
Qy 301 PheTrpLeuGlyGlnIleuValCySlyTrpAlaGlyTrpTrpProTrpAsnIlePhe 320  
Db 964 TTCTGGCTAGAGAGAGCTGTGTGTGTCAGACAGGACCACTCCCTGGAAACATTTTC 1023  
Qy 321 ProValIleSerLeuTrpLeuMetGlyValIleValIleAsnIleSerPheArgIleThrIle 340  
Db 1024 CAGTCACTTCACCTTACCTAATGGGTGAGTTACCAACAGATCTTCGACATCCATC 1083  
Qy 341 LeuProGlnIlnTrpLeuArgProValGluAspValAlaThrSerGlnAspAspCySly 360  
Db 1084 CTTCGCGACATACCTTCGCGGCGAGTGAAGATGGCCACATCCCAAGACATGTTAC 1143  
Qy 361 LysPheAlaIleSerGlnSerSerTrpTrpValIleMetGlyAlaValIleMetGly 380  
Db 1144 AAGTTGGCATCTCACAGTCATCCAGGCACTGTTATGGAGCTGTTATTCATGAGGAC 1203  
Qy 381 PheTrpValValPheAspArgAlaArgLysArgIleGlyPheAlaValSerAlaCySly 400  
Db 1204 TTCTACGTTGTCTTGATGGGCGCCGAAACAAATGGCTTGTGTCAGGCTTGCAT 1263  
Qy 401 ValIleAspGluPheArgTrpAlaAlaValGluGlyProPheValThrLeuAspMetGlu 420  
Db 1264 GTGACAGATGAGTTCAGGACGCGAGGCTGAAGGCTTTTGTACCTTGAACATGGAA 1323  
Qy 421 AspCyGlyTrpAsnIleProGlnThrAspGlnSerThr 433  
Db 1324 GACTGTGGCTACAACTTCCAGACAGATGAGTCAACC 1362

RESULT 4  
US-09-548-367D-3  
; Sequence 3, Application US/09548367D  
; Patent No. 6440698  
; GENERAL INFORMATION:  
; APPLICANT: GURNEY ET AL.  
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREOF AND USES  
; FILE REFERENCE: 29915/6280H  
; CURRENT APPLICATION NUMBER: US/09/548,367D

US-10-726-967a-84 (1-433) x US-09-548-367D-3 (1-2070)

Alignment Scores:  
Pred. No.: 4,34e-257 Length: 2070  
Score: 2267.00 Matches: 429  
Percent Similarity: 99.31% Conservative: 1  
Best Local Similarity: 99.08% Mismatches: 3  
Query Match: 99.08% Indels: 0  
DB: 3 Gaps: 0

US-10-726-967a-84 (1-433) x US-09-548-367D-3 (1-2070)

Qy 1 ThrGlnIleGlyIleArgLeuProLeuArgSerGlyLeuGlyValAlaProLeuGlyLeu 20  
Db 64 ACCAGACAGGACATCCGCTGCTGCTGCGAGGCGCTGGGGGCCCCCTGGGGCTG 123  
Qy 21 GluIleAsnLeuGluTrpAspGluGluProGluGluProGlyIleArgGlyIleSerPhe 40  
Db 124 CGGCTGCCCGGAGAGACCGAGAGAGCCGAGAGCCCGGCGAGGCGGAGGCTTGTG 183  
Qy 41 GluMetValAspAsnLeuArgGlyIySlySerGlyGlnGlyTrpTrpValGluMetThrVal 60  
Db 184 GAGATGTGTGACAACCTGAGGAGGCAAGTGGGGAGGCTTACGTGAGATGACCGTG 243  
Qy 61 GlySerProProGlnThrLeuAsnIleLeuValAspTrpGlySerSerAsnPheAlaVal 80  
Db 244 GGGAGCCCCCGGAGAGGCTCAACATCTGTGTGATACAGGACAGTAACTTTCAGTG 303  
Qy 81 GlyAlaAlaProIleAspPheLeuHisArgTrpTrpGlnArgGlnLeuSerSerThrTrp 100  
Db 304 GGTGTGCCCCCAGACCTTCTGCACTGCTACAGAGGAGCACTGTCCAGCATATC 363  
Qy 101 ArgAspLeuArgLysGlyValIyTrpValProTrpTrpGlnGlyIySerTrpGluGly 120  
Db 364 CGGACCTCCGGAAGGTTGTATGTGCCCTTACACCGAGGCAAGTGGGAAGGAGCTG 423  
Qy 121 GlyThrAspLeuValSerIleProIleProIleValIleValAlaAsnIleAla 140  
Db 424 GGACCGGACCTGTAAAGATCCCATGAGCCCAACGTCATGTGCGGCCAAACATTCCT 483  
Qy 141 AAlIeThrGluSerAspLysPhePheIleAengIySerAsnTrpGluGlyIleuGly 160  
Db 484 GCCATCACTGAATCAAGACAAAGTTCTTCATCAACGGCTCCAACTGGGAAGGATCCTGGG 543  
Qy 161 LeuAlaTrpAlaGluIleAlaArgProAspAspSerIleuGluProPhePheAspSerIleu 180  
Db 544 CTGGCCTATGCTGAATTCAGGAGCTTGAACGATCCGCTCGAGAGCTTCCTTGTGACCTCTG 603  
Qy 181 ValIyGlnThrHisValProAsnLeuPheSerLeuGlnLeuCySGlyAlaGlyPhePro 200  
Db 604 GTAAAGCAGACCCAGTCCCAACCTCTTCCTCCGAGCCTTGTGGTGGCTGGCTTCCC 663  
Qy 201 LeuAengIlnSerGluValIleuAlaSerValGlyIySerMetIleIleGlyIyIleAsp 220  
Db 664 CTCACACAGCTGTAAGTCTGGCTCTGTGAGAGGAGCATGATCATTTGGAGGTATCGAC 723  
Qy 221 HisSerLeuTrpTrpGlySerLeuTrpTrpTrpProIleArgArgGluTrpTrpTrpGlu 240



Db 724 CACTGCTTACACAGGCACTCTCTGTATACACCATCCGGCGGAGTGTATTATGAG 783  
Qy 241 ValIleIleValArgValGluIleAngIyGlnAspLeuLysMetAspCysAlaYsgIuYr 260  
Db 784 GTCATCATTTGTCGGGTGAGATCAATGACAGGATCTGAATAATGACATGCAAGGATAC 843  
Qy 261 AenTYrAspLysSerIleValAspSerGlyThrThrAsnLeuArgLeuProLysIleVal 280  
Db 844 AACATGACAAAGCAATTTGTGACAGTGCACCAACCACTTGTGGTCCCAAGAAAGTG 903  
Qy 281 PheGluAlaIleValLysSerIleValAspSerGlyThrThrAsnLeuArgLeuProLysIleVal 300  
Db 904 TTTGAAGCTGCACTCAATTCATCAAGGACCTCTTCCACGAGAAAGTTCTTATGCT 963  
Qy 301 PheTrpLeuGlyGluGlnLeuValCysTrpGlnAlaGlyThrThrProTrpAsnIlePhe 320  
Db 964 TTTGGCTAGAGAGAGAGCTGGGTGCTGTGCAAGGACGACCACTTCCGAAATTTTC 1023  
Qy 321 ProValIleSerLeuTyrlleuMetGlyGluValThrAsnGlnSerPheArgIleThrIle 340  
Db 1024 CCAATCATCTCACTTAATGGGTGAGTTACCAACCACTTCCGCAATTCAC 1083  
Qy 341 LeuProGlnGlnTyrlleuArgProValGluAspValAlaThrSerGlnAspAspCysTrp 360  
Db 1084 CTTCGCAACATACCTTGGCGGCGAGTGAAGATGTGGCCAGTCCCAAGCACTTTAC 1143  
Qy 361 LysPheAlaIleSerGlnSerSerThrArgIleThrValMetGlyAlaValIleMetGluGly 380  
Db 1144 AGTTTGCATCTCAAGTCATCAAGGAGCTGTTATGGAGCTGTTATCAAGAGGCG 1203  
Qy 381 PheTYrValIlePheAspArgAlaArgIleArgIleGlyPheAlaValSerAlaCysHis 400  
Db 1204 TTTCTAGTTGTCTTTCATCGGGCCGCAAAACGAATGGCTTGTCTTCCAGCGCTTCCAT 1263  
Qy 401 ValHisAspGluPheArgThrAlaAlaValGluGlyProPheValThrLeuAspMetGlu 420  
Db 1264 GTGCAAGATGATTCAGAGAGGCGAGCGGTGGAGAGGCCCTTTTGTCTTGAATGAA 1323  
Qy 421 AspCysGlyTYrAsnIleProGlnThrAspGluSerThr 433  
Db 1324 GACTGTGGCTACAACTTCACAGACAGATGATCAACC 1362

RESULT 5  
US-09-551-853D-3  
Sequence 3, Application US/09551853D  
Patent No. 6500667  
GENERAL INFORMATION:  
APPLICANT: GURNEY ET AL.  
TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES  
FILE REFERENCE: 29915/6280L  
CURRENT APPLICATION NUMBER: US/09/551,853D  
CURRENT FILING DATE: 2000-04-18  
PRIOR APPLICATION NUMBER: US 60/155,493  
PRIOR FILING DATE: 1999-09-23  
PRIOR APPLICATION NUMBER: US 09/404,133  
PRIOR FILING DATE: 1999-09-23  
PRIOR APPLICATION NUMBER: PCT/US99/20881  
PRIOR FILING DATE: 1999-09-23  
PRIOR APPLICATION NUMBER: US 60/101,594  
PRIOR FILING DATE: 1998-09-24  
NUMBER OF SEQ ID NOS: 73  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 3  
LENGTH: 2070  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-551-853D-3

Alignment Scores:  
Pred. No.: 4,34e-257 Length: 2070  
Score: 2267.00 Matches: 429  
Percent Similarity: 99.31% Conservative: 1

Best Local Similarity: 99.08% Mismatches: 3  
Query Match: 99.08% Indels: 0  
DB: 4 Gaps: 0  
US-10-726-967a-84 (1-433) x US-09-551-853D-3 (1-2070)

Qy 1 ThrGlnHisGlyIleArgLeuProLeuArgSerGlyLeuGlyAlaProLeuGlyLeu 20  
Db 64 ACCAGCACCGGATCCGCTGCTCCCTGGCAGCGGCTCGGGGGGGCGCCCTGGGGCTG 123  
Qy 21 GluIleAsnLeuGluThrAspGluGluProGluGluProGlyAlaArgArgIleSerPheVal 40  
Db 124 CGGCTGCCCGGAGAACCGAGAGAGCCCGAGAGAGCCCGGAGGAGGCGAGCTTTGTG 183  
Qy 41 GluMetValAspAsnLeuArgGlyLysSerGlyGlnGlyTYrTYrValGluMetThrVal 60  
Db 184 GAGATGTGTGACAACTTGAAGGGGCAAGTGGGGCGGGCTCTACTAGTGAAGTACCGTG 243  
Qy 61 GlySerProProGlnThrIleuAsnIleuValAspThrGlySerSerAsnPheAlaVal 80  
Db 244 GGCAGCCCCCGCAGACCGCTCAACATCTGTGTGATACAGGCAAGATTACTTTCAGTG 303  
Qy 81 GlyAlaAlaProHisProPheLeuHisArgTYrTYrGlnArgGlnLeuSerSerThrTYr 100  
Db 304 GGTGTGCCCCCACCCTTCTGTCATGCTACAGAGGCACTGTCCAGCAATAC 363  
Qy 101 ArgAspLeuArgLysGlyValTYrValProTYrThrGlnGlyLysTrpGluGluLeu 120  
Db 364 CGGAGCTTCGGAAAGGTGTGTATGTGCTTACACCCAGGCGCAAGTGGAGAGGAGCTG 423  
Qy 121 GlyThrAspLeuValSerIleProHisGlyProAsnValThrValArgAlaAsnIleAla 140  
Db 424 GGCACCGACCTGTGTAAGCATCCCAAGGCCCAAGTCACTGTGCTGCCAATCTGCT 483  
Qy 141 AlaIleThrGlnSerAspLysPhePheIleAsnGlySerAsnTrpGluGlyIleLeuGly 160  
Db 484 GGCATCATCGATGACAGAACTTTCATCAACGCGCTCCAATGGGAAGGCACTCTGGGG 543  
Qy 161 LeuAlaTYrAlaGluIleAlaArgProAspAspSerLeuGluProPheAspSerLeu 180  
Db 544 CTGGCTTATGCTGAATTCGAGGCTGACGACTCCCTCGAGACCTTTTCTTAACCTCTG 603  
Qy 181 ValLysGlnThrHisValProAsnLeuPheSerLeuGlnLeuCysGlyAlaGlyPhePro 200  
Db 604 GTTAAAGCAAGCCAGATTCCAACTCTTCTTCTCGACCTTGTGGTCTGCTTCC 663  
Qy 201 LeuAsnGlnSerGluValLeuAlaSerValGlyLysMetIleIleGlyIleAsp 220  
Db 664 CTCAACCACTCGAAGTGTGCTGCTGTGCGAGGAGCATGATCATTTGAGAGTATCGAC 723  
Qy 221 HisSerLeuTYrThrGlySerLeuTrpTYrThrProIleArgArgGluTrpTYrTYrGlu 240  
Db 724 CACTGCTGTACACAGGAGCTCTCTGTATACACCATCCGCGGAGGTGATTAATGAG 783  
Qy 241 ValIleIleValArgValGluIleAngIyGlnAspLeuLysMetAspCysAlaYsgIuYr 260  
Db 784 GTCATCATTTGTCGGGTGAGATCAATGACAGGATCTGAATAATGACATGCAAGGATAC 843  
Qy 261 AenTYrAspLysSerIleValAspSerGlyThrThrAsnLeuArgLeuProLysIleVal 280  
Db 844 AACATGACAAAGCAATTTGTGACAGTGCACCAACCACTTGTGGTCCCAAGAAAGTG 903  
Qy 281 PheGluAlaIleValLysSerIleValAspSerGlyThrThrAsnLeuArgLeuProLysIleVal 300  
Db 904 TTTGAAGCTGCACTCAATTCATCAAGGACCTCTTCCACGAGAAAGTTCTTATGCT 963  
Qy 301 PheTrpLeuGlyGluGlnLeuValCysTrpGlnAlaGlyThrThrProTrpAsnIlePhe 320  
Db 964 TTTGGCTAGAGAGAGAGCTGGGTGCTGTGCAAGGACGACCACTTCCGAAATTTTC 1023  
Qy 321 ProValIleSerLeuTyrlleuMetGlyGluValThrAsnGlnSerPheArgIleThrIle 340  
Db 1024 CCAATCATCTCACTTAATGGGTGAGTTACCAACCACTTCCGCAATTCAC 1083

QY 341 LeuProGlnGlnTyrLeuAspArgProValGluAspValAlaThrSerGlnAspAspCysTyr 360  
Db 1084 CTTCCGACGAAATACCTGGCGGCGAAGATGTGGCCACGTCACCAAGACGATGTTAC 1143  
QY 361 LysPheAlaIleSerGlnSerSerThrGlyThrValMetGlyValIleMetGlnGly 380  
Db 1144 AAGTTGGCATCTCAAGTCATCCACGGGCACTGTTAGAGAGCTGTATCATGAGGGC 1203  
QY 381 PheTyrValValPheAspArgAlaArgGlyAspGllleGlyPheAlaValSerAlaCysHis 400  
Db 1204 TTCTACGTTGCTTGTGATCGGGCCGCAAAACGAATGGCTTGTGCTGTCAGCGCTTGCAT 1263  
QY 401 ValHisAspGluPheArgThrAlaAlaValGluGlyProPheValThrLeuAspMetGlu 420  
Db 1264 GTGACGAGTACGATTCAGACGCGACCGGTGAAGGCCCTTTGTCACTTGACATGGAA 1323  
QY 421 AspCysGlyTyrAsnIleProGlnThrAspGluSerThr 433  
Db 1324 GACTGTGGCTACACATTCACAGACAGATGATCAACC 1362

## RESULT 6

US-09-416-901B-3  
; Sequence 3, Application US/09416901B  
; Patent No. 6699671  
; GENERAL INFORMATION:  
; APPLICANT: GURNEY ET AL.  
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES  
; FILE REFERENCE: 29915/6280A  
; CURRENT APPLICATION NUMBER: US/09/416,901B  
; PRIOR FILING DATE: 1999-10-13  
; PRIOR APPLICATION NUMBER: US 60/155,493  
; PRIOR FILING DATE: 1999-09-23  
; PRIOR APPLICATION NUMBER: US 09/404,133  
; PRIOR FILING DATE: 1999-09-23  
; PRIOR APPLICATION NUMBER: PCT/US99/20881  
; PRIOR FILING DATE: 1999-09-23  
; PRIOR APPLICATION NUMBER: US 60/101,594  
; NUMBER OF SEQ ID NOS: 72  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 2070  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-416-901B-3

## Alignment Scores:

Pred. No.: 4,34e-257 Length: 2070  
Score: 2267.00 Matches: 429  
Percent Similarity: 99.31% Conservative: 1  
Best Local Similarity: 99.08% Mismatches: 3  
Query Match: 99.08% Indels: 0  
DB: 4 Gaps: 0

US-10-726-967A-84 (1-433) x US-09-416-901B-3 (1-2070)

QY 1 ThrGlnHisGlyIleArgLeuProLeuArgSerGlyLeuGlyGlyAlaProLeuGlyLeu 20  
Db 64 ACCCGACAGGACATCGGCTGCCCTGGCGACGGCCCTGGGGGGGCCCCCTGGGGCTG 123  
QY 21 GlnIleAsnLeuGluThrAspGluGluProGluGluProGlyArgArgGlySerPheVal 40  
Db 124 CGGCTGCCCGGAGACCGACGAAAGCCGACGAGGCCGCCGGAAGGGGAGCTTTGTG 183  
QY 41 GluMetValAspAsnLeuArgGlyLysSerGlyGlnGlyTyrTyrValGluMetThrVal 60  
Db 184 GAGATGGTGGACAACTTGAAGGGGCAAGTCGGGGCAGGCTACTACGTGAGATGACCGTG 243  
QY 61 GlySerProGlnThrLeuAsnIleValAspThrGlySerSerAsnPheAlaVal 80  
Db 244 GGCAGCCCCCGGACGACCTCAATCTGTGTGATACAGGACAGTAACCTTGACGTG 303

QY 81 GlyAlaAlaProHisProPheLeuHisArgTyrTyrGlnAspGlnLeuSerSerThrTyr 100  
Db 304 GGTGCTGCCCCCACCCTCTTCTGTCATCGTACCTACCGAAGAGCAGCTGTCCAGCACATAC 363  
QY 101 ArgAspLeuArgLysGlyValTyrValProTyrThrGlnGlyLysTPGlnGlyLeu 120  
Db 364 CGGACCTCCGAAAGGGTGTATGTCGCCCTACACCGGCGCAAGTGGAAAGGGGAGCTG 423  
QY 121 GlyThrAspLeuValSerIleProHisGlyProAsnValThrValArgAlaAsnIleAla 140  
Db 424 GGCACCGACCTGGTAAAGCATCCCCCATGCCCCCAACGTCACTGTGCGTCAACATTCCT 483  
QY 141 AlaIleThrGlnSerAspLysPhePheIleAsnIleYerAsnTPGlnGlyIleLeuGly 160  
Db 484 GCCATCATGAAATCAGACAAAGTTTTCATCAACGGCTCCAACTGGGAAGGACATCCGGGG 543  
QY 161 LeuAlaTyrAlaGluIleAlaArgProAspAspSerLeuGluProPhePheAspSerLeu 180  
Db 544 CTGGCTATGTCTGAGATTGCGAGGCTGACGACTGCCCTGAGGCTTTCTTGACTCTTG 603  
QY 181 ValLysGlnThrHisValProAsnLeuPheSerLeuGlnLeuCysGlyAlaGlyPhePro 200  
Db 604 GTAAAGCAGACCCAGCTTCCAACTCTTCTCCCTGACGCTTGTGTGTGCTGCTTCCCC 663  
QY 201 LeuAsnGlnSerGluValLeuAlaSerValGlySerMetIleIleGlyGlyIleAsp 220  
Db 664 CTCAACAGACTGAAGTGTGGCTGCTGTGCGAGGAGATATCATTTGAGGTATCGAC 723  
QY 221 HisSerLeuTyrThrGlySerLeuTyrTyrThrProIleArgArgGluTyrTyrGlu 240  
Db 724 CACTCGCTGTACACAGGACGATCTCTGTATACACCATCCGCGGGAGGTGATATATGAG 783  
QY 241 ValIleIleValAlaArgValGluIleAsnGlyGlnAspLeuLysMetAspCysLysGlyTyr 260  
Db 784 GTCATCATGTGCGGGGAGATCATGACAGAGATCTAAATGTGATCGCAAGGAGTAC 843  
QY 261 AsnTyrAspLysSerIleValAspSerGlyThrThrAsnLeuArgLeuProLysVal 280  
Db 844 AACTATGACAAAGGACATTTGTGACATGTGCACACCAACTTCGTTGCCCAAGAAAGTG 903  
QY 281 PheGluAlaAlaValLysSerIleLysAlaIleSerSerThrGluLysPheProAspGly 300  
Db 904 TTTGAAGCTGAGTCAATCATCATCAAGGACGCTCTTCCAGGAAAGTCCCTGATGT 963  
QY 301 PheThrLeuGlyGluGlnLeuValCysTrpGlnAlaGlyThrThrProTrpAsnIlePhe 320  
Db 964 TTCTGGCTAGGAGAGAGCTGTGTGTGCGCAAGCAGCACCACTTGGTGAACATTTTC 1023  
QY 321 ProValIleSerLeuTyrLeuMetGlyGluValThrAsnGlnSerPheArgIleThrIle 340  
Db 1024 CCAGTATCTCACTTCAATTAAGGTGAGTTCACAAACGATCTTCCGATCACATC 1083  
QY 341 LeuProGlnGlnTyrLeuArgProValGluAspValAlaThrSerGlnAspAspCysTyr 360  
Db 1084 CTTCCGACGAAATACCTCGCGCAGTGAAGATGGGCAAGTCCCAAGACGATGTAC 1143  
QY 361 LysPheAlaIleSerGlnSerSerThrGlyThrValMetGlyAlaValIleMetGlnGly 380  
Db 1144 AAGTTGGCATCTCAAGTCATCCACGGGCACTGTTAGAGAGCTGTATCATGAGGGC 1203  
QY 381 PheTyrValValPheAspArgAlaArgGlyAspGllleGlyPheAlaValSerAlaCysHis 400  
Db 1204 TTCTACGTTGCTTGTGATCGGGCCGCAAAACGAATGGCTTGTGCTGTCAGCGCTTGCAT 1263  
QY 401 ValHisAspGluPheArgThrAlaAlaValGluGlyProPheValThrLeuAspMetGlu 420  
Db 1264 GTGACGAGTACGATTCAGACGCGACCGGTGAAGGCCCTTTGTCACTTGACATGGAA 1323  
QY 421 AspCysGlyTyrAsnIleProGlnThrAspGluSerThr 433  
Db 1324 GACTGTGGCTACACATTCACAGACAGATGATCAACC 1362

```
RESULT 7
US-09-548-376D-3
; Sequence 3, Application US/09548376D
; Patent No. 6706485
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR
; TITLE OF INVENTION: AND USES
; FILE REFERENCE: 29915/6280F
; CURRENT APPLICATION NUMBER: US/09/548,376D
; CURRENT FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 2070
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-548-376D-3

Alignment Scores:
Pred. No.: 4,34e-257 Length: 2070
Score: 2267.00 Matches: 429
Percent Similarity: 99.31% Conservative: 1
Best Local Similarity: 99.08% Mismatches: 3
Query Match: 99.08% Indels: 0
Gaps: 0

US-10-726-967A-84 (1-433) x US-09-548-376D-3 (1-2070)
QY 1 ThrGlnHieGlyIleArgLeuProLeuArgSerGlyLeuGlyValAlaProLeuGlyLeu 20
DB 64 ACCCAACACGGGATCCGGCTGGCCCTGGCGAGCGGCTGGGGGGCCCCCTGGGGGCTG 123
QY 21 GlnIleAsnLeuGlyThrArgGlyGluProGlyGluProGlyValArgGlySerPheVal 40
DB 124 CGGCTCCCGGGGAGACCGGACGAGGAGCCCGGAGGAGCCCGGAGGCGGACCTTGTG 183
QY 41 GluMetValAspMetLeuArgGlyLysSerGlyGlnGlyTyrTyrValGluMetThrVal 60
DB 184 GAGATGGTGGCAACCTGAGGGGCAAGTCGGGGCAAGGCTACTACGTGGAGTACCCTG 243
QY 61 GlySerProGlnThrLeuAsnIleLeuValAspThrGlySerSerAsnPheAlaVal 80
DB 244 GGGAGCCCCCGGAGAGCTCAACATCTGTGTGATACAGGACGACTTGTGACGTG 303
QY 81 GlyAlaAlaProHisProPheLeuHisArgTyrTyrGlnArgGlnLeuSerSerThrTyr 100
DB 304 GGTGTGGCCCCCAACCCCTTCTGACATCGCTACTACACAGGACGCTGCACACATAC 363
QY 101 ArgAspLeuArgGlyValTyrValProTyrThrGlnGlyLysTyrGluGluLeu 120
DB 364 CGGAGCTTCGGAGAGGTGTGTATGTGCTCTACACCCAGGGCAAGTGGAGAGGAGCTG 423
QY 121 GlyThrAspLeuValSerIleProHisGlyProAsnValThrValArgAlaAsnIleAla 140
DB 424 GGCACGCACTGTGTAGCATCCGCCATGGCCCAACTCACTGTGGCTGCAACATTCCT 483
QY 141 AlaIleThrGlySerAspLysPhePheIleAsnGlySerAsnTyrGluGlyIleLeuGly 160
DB 484 GCCATCACTGAATCAGACAGTCTTCTCATCAACGGCTCCAACTGGGAGGACATCTGGGG 543
QY 161 LeuAlaTyrTrpAlaGluIleAlaArgProAspAspSerLeuGluProPheAspSerLeu 180
DB 544 CTGGCTATGCTGAGATTGGCCAGGCTGACGACTCCCTGGAGCCTTCTTGTGACTCTG 603
```

```
QY 181 ValIleGlyIleThrHisValProAsnLeuPheSerLeuGlnLeuCyseGlyValAlaGlyPhePro 200
DB 604 GTTAAAGACACCCAGGATCCCAACTCTTCTCCAGCTTTGTGTGCTGCTGCTTCC 663
QY 201 LeuAsnGlnSerGlyValLeuAlaSerValGlyGlySerMetIleIleGlyIleAsp 220
DB 664 CTCACCACTGTGAAGTGTCTGCTCTGTGGAGGAGCATGATCATTTGAGATTCGAC 723
QY 221 HisSerLeuTyrThrGlySerLeuTyrThrProIleArgArgGluTyrTyrGlu 240
DB 724 CACTCGCTGTACACAGGAGTCTGTGTATACACCACTCCGCGGAGGTGATATATAG 783
QY 241 ValIleIleValArgValGluIleAsnGlyGlnAspLeuLysMetAspCyseGlyTyr 260
DB 784 GTCATCATTTGGCGGGGTGAGTCAATGACAGATCTGAAATGACTGACAGAGTAC 843
QY 261 AsnTyrAspLysSerIleValAspSerGlyThrThrAsnLeuArgLeuProLysVal 280
DB 844 AACTTGTACAGAGCATTTGTGACAGTGGACACCAACCTTGTGGCCAAAGTGTG 903
QY 281 PheGlnAlaAlaValLysSerIleLysAlaAlaSerSerThrGluLysPheProAspGly 300
DB 904 TTTGAAGCTGACGTCAAATCCATCAAGGACGCTCTCCACGAGAAAGTTCCTGATGT 963
QY 301 PheTyrPheGlyGlyGlnLeuValCysTyrGlnAlaGlyThrThrProThrAsnIlePhe 320
DB 964 TTCTGGCTAGGAGACAGCTGTGTGTGGCGAGGACGACCAACCCCTTGGAACTTTTC 1023
QY 321 ProValIleSerLeuTyrLeuMetGlyGlnValThrAsnGlnSerPheArgIleThrIle 340
DB 1024 CAGTCATCTTCACTTACTTAACTGAGTGAAGTTACCAACCAATCTCTCCGATCACCATC 1083
QY 341 LeuProGlnGlnTyrLeuArgProValGluAspValAlaThrSerGlnAspAspCysTyr 360
DB 1084 CTTCCGACAGATATACCTGGCGCCAGTGGAAAGTGTGCCACCTCCCAAGACACATGTTTC 1143
QY 361 LysPheAlaIleSerGlnSerSerThrArgTyrThrValMetGlyAlaValIleMetGlyGly 380
DB 1144 AAGTTTGCATCTTCAAGTATCCAGGACGCTGTATAGGAGCTGTATCATGAGGGC 1203
QY 381 PheTyrValIlePheAspArgAlaArgLysArgIleGlyPheAlaValSerAlaCysHis 400
DB 1204 TTCTACGTTGTCTTGTATGTCGGCCGGAACCAATTTGGCTTGTGCTGACGCTGCCAT 1263
QY 401 ValHisAspGluPheArgThrAlaAlaValGluGlyProPheValThrLeuAspMetGlu 420
DB 1264 GTGACAGATGAGTTCAGACGCGCAGGGGTGAAAGCCCTTTTGTCACTTGGACATGAA 1323
QY 421 AspCyseGlyTyrAsnIleProGlnThrAspGlySerThr 433
DB 1324 GACTGTGCTACAAATTCACAGACAGATGATCAACC 1362

RESULT 8
US-09-794-927A-3
; Sequence 3, Application US/09794927A
; Patent No. 6727074
; GENERAL INFORMATION:
; APPLICANT: Gurney et al.
; TITLE OF INVENTION: Alzheimer's Disease Secretase, APP Substrates Therefor and Uses
; TITLE OF INVENTION: Therefor
; FILE REFERENCE: 29915/6280FG
; CURRENT APPLICATION NUMBER: US/09/794,927A
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 09/416,901
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101,594
```

PRIOR FILING DATE: 1998-09-24  
NUMBER OF SEQ ID NOS: 74  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 3  
LENGTH: 2070  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-794-927A-3

Alignment Scores:

Pred. No.:	4,346-257	Length:	2070
Score:	2267.00	Matches:	429
Percent Similarity:	99.31%	Conservative:	3
Best Local Similarity:	99.08%	Mismatches:	1
Query Match:	99.08%	Indels:	0
DB:	4	Gaps:	0

US-10-726-967A-84 (1-433) x US-09-794-927A-3 (1-2070)

QY 1 ThrGlnHisGlyIleArgLeuProLeuArgSerGlyLeuGlyValAlaProLeuGlyLeu 20  
Db ACCACGACGCGCATCCGGCTCCCTCGACAGCGGCTCGGGGGGCGCCCTCGGGGCTG 123  
QY 21 GluIleAsnLeuGluThrAspGluGluProGluGluArgArgGlySerPheVal 40  
Db CGGCTGCCCGGAGACCGACGAGAGACCGGAGGAGCCGGCCGGAGGGGCACTTTGTC 183  
QY 41 GluMetValAspAsnLeuArgGlyLysSerGlyGlnGlyTyrTyrValGluMetThrVal 60  
Db GAGATGGTGAACAACTGAGGGGCAAGTCGGGGCAAGGCTCACTGAGATGACCGTG 243  
QY 61 GlySerProGlnThrLeuAsnIleLeuValAspThrGlySerSerAsnPheAlaVal 80  
Db GGCAGCCCCCGGACGATCCATCATCTGCTGATGATGACGACGATGATCTTTCAGTTC 303  
QY 81 GlyAlaIleAlaProHisProPheLeuHisArgTyrTyrGlnArgGlnLeuSerThrTyr 100  
Db GGATGCTCCCCCAACCTTCTTCGACATCGCTACTACAGAGGACGCTGCCAGACATAC 363  
QY 101 ArgAspLeuArgLysGlyValTyrValProTyrThrGlnGlyLysThrGluGluLeu 120  
Db CGGACCTCCGGAAGGCTGTGTGTGCTTACACCCAGGCAAGTGGGAAAGGGAGCTG 423  
QY 121 GlyThrAspLeuValSerIleProHisGlyProAsnValThrValArgAlaAsnIleAla 140  
Db GGCACGACCTGGTGAAGCATCCGCCATGCCCCCAAGCTGACGTGCTCCAACTTCT 483  
QY 141 AlaIleThrGlySerAspLysPhePheIleAsnGlySerAsnThrGluGlyIleLeuGly 160  
Db GGCATCATCGAATCACAACAAGTTCTTCACTAACGGCTCCACTGGAAAGGCACTTGGGG 543  
QY 161 LeuAlaTyrAlaGluIleAlaArgProAspAspSerLeuGluProPhePheAspSerLeu 180  
Db CTGGCTTATGCTGAGATTCGCCAGGCTGACGACTCCCTGAGACTTTCTTTGACTCTTC 603  
QY 181 ValIlyGlnThrHisValProLeuPheSerLeuGlnLeuGlyAlaGlyPhePro 200  
Db GTHAAGACAGCCCAAGTCCCAACTTCTTCCGACGCTTGTGGTGGCTGCTCC 663  
QY 201 LeuAsnGlnSerGluValLeuAlaSerValGlyGlySerMetIleIleGlyIleAsp 220  
Db CTCAACCAAGTCTGAGTGTGGCTCTCTCGAGGAGACATGATTCATTGGAGATTCGAC 723  
QY 221 HisSerLeuTyrThrGlySerLeuTyrTyrThrProIleArgArgGluTyrTyrGlu 240  
Db CACTGCTGATACAGGAGTCTCTGATATACACCACTCCGCGGAGTGTATTATGAG 763  
QY 241 ValIleIleValArgValGluIleAsnGlyGlnAspLeuLysMetAspCysLysGluTyr 260  
Db GTCATCATTTGGCGGCTGAGATCATGAGACAGATCTGAAATATGACTGCAAGAGATAC 843  
QY 261 AsnTyrAspLysSerIleValAspSerGlyThrThrAsnLeuArgLeuProLysVal 280

Db 844 AACTATGACAGACATTTGTGACAGTGGACACCAACTTCTGTTGCCAAGAAAGTG 903  
QY 281 PheGluAlaIleValLysSerIleLysAlaIleAspSerThrGluLysPheProAspGly 300  
Db TTTGAAGCTGACAGTCAATTCATTCAGAGGACACTCTCCACGAGAAAGTTCCCTGATGGT 963  
QY 301 PheTyrPheGluGlyGluIleLeuValCysThrGlnIleArgIleThrProTyrPheIlePhe 320  
Db TTCTGGCTTGAAGACAGCACTGTGTGTGCGAAGGACGACACCCCTGGAAACATTTTC 1023  
QY 321 ProValIleSerLeuTyrLeuMetGlyGlyValThrAsnGlnSerPheArgIleThrIle 340  
Db CCAGTCATCTCACTCACTCACTAATGGGTAGGTTTCAACACAGTCTCTCCGATCACCATC 1083  
QY 341 LeuProGlnGlnTyrLeuAsnGlnProValGluAspValAlaThrSerGlnAspAspCysTyr 360  
Db CTTCCGACGAAATACCTGGGCGCAGTGAAGATGTGCCACGTTCCCAAGACGACTGTAC 1143  
QY 361 LysPheAlaIleSerGlnSerSerThrGlyThrValMetGlyAlaValIleMetGluGly 380  
Db AAGTTGCCATCTCAAGTATCCACGGGCACTTTATGGAGCTGTTATCATGAGAGGC 1203  
QY 381 PheTyrValValPheAspArgAlaArgLysArgIleGlyPheAlaValSerAlaCysHis 400  
Db TTCTACGTTGCTTGTGATCGGGCCCGGAAACGAATGGCTTTCGTGTCAGCGCTTGCAT 1263  
QY 401 ValHisAspGluPheArgThrAlaAlaValGluGlyProPheValThrLeuAspMetGlu 420  
Db GTGCACGATGATGTTCAAGACGCGACGCGGTGGAAGGCCCTTTGTCTACCTTGGACATGAA 1323  
QY 421 AspCysGlyTyrAsnIleProGlnThrAspGlnSerThr 433  
Db GACTGTGGCTTAAACAATTCACAGACAGATGATCAACC 1362

RESULT 9

US-09-548-373D-3  
Sequence 3, Application US/09548373D  
Patent No. 6737510  
GENERAL INFORMATION:  
APPLICANT: GURNEY ET AL.  
TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES  
FILE REFERENCE: 29915/62808  
CURRENT APPLICATION NUMBER: US/09/548,373D  
PRIOR FILING DATE: 2000-04-12  
PRIOR APPLICATION NUMBER: US 60/155,493  
PRIOR FILING DATE: 1999-09-23  
PRIOR APPLICATION NUMBER: US 09/404,133  
PRIOR FILING DATE: 1999-09-23  
PRIOR APPLICATION NUMBER: PCT/US99/20881  
PRIOR FILING DATE: 1999-09-23  
PRIOR APPLICATION NUMBER: US 60/101,594  
PRIOR FILING DATE: 1998-09-24  
NUMBER OF SEQ ID NOS: 73  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 3  
LENGTH: 2070  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-548-373D-3

Alignment Scores:

Pred. No.:	4,346-257	Length:	2070
Score:	2267.00	Matches:	429
Percent Similarity:	99.31%	Conservative:	1
Best Local Similarity:	99.08%	Mismatches:	3
Query Match:	99.08%	Indels:	0
DB:	4	Gaps:	0

US-10-726-967A-84 (1-433) x US-09-548-373D-3 (1-2070)

QY 1 ThrGlnHisGlyIleArgLeuProLeuArgSerGlyLeuGlyValAlaProLeuGlyLeu 20

64 ACCGACACGGGATCCGGCTCCCTCCGACAGCGGCTGGGGGGCCCCCTGGGGCTG 123  
21 GJUIAaenleuGluThraApGluGluProGluGluProGluYArgArgLySerPheVal 40  
124 CGGCTGCCCGGGAGACGAGAGAGGCGAGAGCGGCGGAGGGGGGAGCTTTGG 183  
41 GJUIaerValAaPaaenleuArgLyLySerGlyGlnGlyTYrTYrValGJUIaerThraVal 60  
184 GAGATGTGTGACAACTGAGGGGCAAGTGGGGGCGAGGCTACTAGGTGAGATGACCGG 243  
61 GJYserProProGluThraenleuValAaPaaThraGlySerSerAaenPheAaVal 80  
244 GGCAGCGCCCGCAGACGCTCAACATCTGTGTGATACAGGACAGATTAATTTCAGTG 303  
81 GJYAlaAaPaaProHiaProPheleuHiaArgTYrTYrGlnArgGlnleuSerSerThraTYr 100  
304 GGTGTGCTGCCCGACCCCTTCTGATGCTACTACCAAGGAGGTGTCCAGACATAC 363  
101 ArgAaPaaenArgLyGlyValTYrValProTYrThraGlnGlyLyGTPGluGluVal 120  
364 CGGAGCTTCGGAGAGGGTGTGTATGTGCTTACACCCAGGCGCAAGTGGAGAGGGAGCTG 423  
121 GJYThraApLeuValSerLleProHiaGlyProAaenValThraValArgAlaAaenleuAa 140  
424 GGCACCGACCTGTGTAAAGCATCCCGACGCGCCCAAGGTCACTGTGCTGCCAATTCCT 483  
141 AlaLleThraGlySerAaPaaPhePheLleAaenGlySerAaenTPGluGlyLleuGly 160  
484 GGCATCACTGAAATCAAGACAAATGTTCTTCAATCAACGGGTCCACAGGAAAGCATTCGGGG 543  
161 LeuAaLleTYrAlaGluLleAaArgProAaPaaPaaSerLleuGluProPheAaPaaSerLleu 180  
544 CTGGCTTATGTGAGATGCTCCAGGCTCGACGACTCCCTGAGGCTTTTGTGACTCTCG 603  
181 ValLySerGlnThraHiaValProAaenleuPheSerLleuGluGlyValAaGlyPhePro 200  
604 GTAAGACAGACCCAGCTTCCCAACTTCTCTCCCTCAAGTGTGTGTGCTGCTTCC 663  
201 LeuAaenGlnSerGlyValleuAaSerValGlyGlySerMetLleLleGlyGlyLleAaP 220  
664 CTCACCAATCTCAATGCTGCTGCTGTGAGGAGATGATCATTTGAGATTCAGAC 723  
221 HiaSerLleuTYrThraGlySerLleuTYrThraProLleArgArgGluTYrTYrGlu 240  
724 CACTCGCTGTACACAGGAGTCTGTGTATACACCATCCGCGGAGTGTGTATTTAGAG 783  
241 ValLleLleValArgValGluLleAaenGlyGlnAaPaaLleuPheAaPaaPaaSerLleuTYr 260  
784 GTCATCATTTGTGGGTGTGAGATCAATGACAGATCTGAAATGAGACTGCAAGAGATAC 843  
261 AaenTYrAaPaaLySerLleValAaPaaSerGlyTYrThraAaenleuArgLeuProLyLyVal 280  
844 AACTATGACAAAGCATTTGTGACATGGGACCAACCACTTCGTTTGCACCAAGAAAGTG 963  
281 PheGluAaAaValLySerLleuValAaLleAaSerSerThraGlyLySerPheProAaPaaGly 300  
904 TTTGAAGCTGCACTCAATCAATCAAGGAGCTCTCCCAAGGAGATTCCTGTAGTG 963  
301 PheTPLeuGlyGlyGluLleuValCyETTPGlnAaGlyTYrThraProTPAaenLlePhe 320  
964 TTTGTGCTAGAGAGAGCTGTGTGTGTGCAAGAGGACCAACCTTTGGAACATTTTC 1023  
321 ProValLleSerLleuTYrLleuMetGlyGluValThraAaenLlePheArgLleThraLle 340  
1024 CCACTCATCTCACTCACTCAATGAGGTGAGGTTCACCAACAGCTCTTCGCCATCACATC 1083  
341 LeuProGlnGlnTYrLleuArgProValGluAaPaaValAaThraSerGlnAaPaaPaaCyETTYr 360  
1084 CTTCGAGAGCAATACCTGCGGCGAGTGAAGATGTGGCCACGTCCCAAGACGACTGTAC 1143  
361 LyPhePheAaLleSerGlnSerSerThraGlyThraValMetGlyValAaLleLleMetGlyGly 380  
1144 AAGTTTGCATCTCAACATCAACGAGGACGTGTATGAGGACGTGTATCAATGAGGAGC 1203

QY 381 PheTYrValValPheAaPaaArgAlaArgLyArgGlyLleGlyPheAaValSerAlaCyAHis 400  
DB 1204 TTTCAAGTTGCTTTGATCGGAGCCCGAAACGAATGTGCTTTGCTGTACAGCTTGGCAT 1263  
QY 401 ValHiaAaPaaGluPheArgThraAlaValGluGlyProPheValThraAaPaaMetGlu 420  
DB 1264 GTGCAAGTGAAGTTCAAGACGCGAGCGGTGGAAGGCCCTTTGTCACTTGTGACATGGA 1323  
QY 421 AaPaaGlyTYrAaenLleProGlnThraApGlySerThra 433  
DB 1324 GACTGTGCTCAACATTCACAGACAGATGATGATCAACC 1362  
RESULT 10  
US-09-795-847B-3  
/ Sequence 3, Application US/09795847B  
/ Patent No. 6753163  
/ GENERAL INFORMATION:  
/ APPLICANT: Guiney, Mark E.  
/ APPLICANT: Bienkowski, Michael J.  
/ APPLICANT: Heimlikson, Robert L.  
/ APPLICANT: Parodi, Luis A.  
/ APPLICANT: Yan, Riqiang  
/ TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND USES  
/ FILE REFERENCE: 28341/6280DE  
/ CURRENT APPLICATION NUMBER: US/09/795,847B  
/ PRIOR FILING DATE: 2001-02-28  
/ PRIOR APPLICATION NUMBER: 09/416,901  
/ PRIOR FILING DATE: 1999-10-13  
/ PRIOR APPLICATION NUMBER: 60/155,493  
/ PRIOR FILING DATE: 1999-09-23  
/ PRIOR APPLICATION NUMBER: 09/404,133  
/ PRIOR FILING DATE: 1999-09-23  
/ PRIOR APPLICATION NUMBER: PCT/US99/20881  
/ PRIOR FILING DATE: 1999-09-23  
/ PRIOR APPLICATION NUMBER: 60/101,594  
/ PRIOR FILING DATE: 1998-09-24  
/ NUMBER OF SEQ ID NOS: 74  
/ SOFTWARE: PatentIn Ver. 2.0  
/ SEQ ID NO 3  
/ LENGTH: 2070  
/ TYPE: DNA  
/ ORGANISM: Homo sapiens  
US-09-795-847B-3  
Alignment Scores:  
Pred. No.: 4,34e-257 Length: 2070  
Score: 2267.00 Matches: 429  
Percent Similarity: 99.31% Conservative: 1  
Best Local Similarity: 99.08% Mismatches: 3  
Query Match: 99.08% Indels: 0  
DB: 4 Gaps: 0  
US-10-726-967a-84 (1-433) x US-09-795-847B-3 (1-2070)  
QY 1 ThraGlnHiaGlyLleArgLeuProleuAaGserGlyLleuGlyGlyAlaProleuGlyLeu 20  
DB 64 ACCGACACGGGATCCGGCTCCCTCCGACAGCGGCTGGGGGGGGCCCCCTGGGGCTG 123  
QY 21 GJUIAaenleuGluThraApGluGluProGluGluProGluYArgArgLySerPheVal 40  
DB 124 CGGCTGCCCGGGAGACGAGAGAGGCGAGAGCGGCGGAGGGGGGAGCTTTGG 183  
QY 41 GJUIaerValAaPaaenleuArgLyLySerGlyGlnGlyTYrTYrValGJUIaerThraVal 60  
DB 184 GAGATGTGTGACAACTGAGGGGCAAGTGGGGGCGAGGCTACTAGGTGAGATGACCGG 243  
QY 61 GJYserProProGlnThraenleuValAaPaaThraGlySerSerAaenPheAaVal 80  
DB 244 GGCAGCGCCCGCAGACGCTCAACATCTGTGTGATACAGGACAGATTAATTTCAGTG 303  
QY 81 GJYAlaAaPaaProHiaProPheleuHiaArgTYrTYrGlnArgGlnleuSerSerThraTYr 100

Db 304 GGTCTGCCCCCACCCTCTCTCATCGCTACTACAGAGGCGAGCTGTCCAGACATAC 363  
QY 101 ArgAspLeuArglysglyValTyrValProTyrThrGlnGlyLysTyrPGLuGlyGluLeu 120  
Db 364 CGGAGACCTCCGGAGGGGTGTATGTGCTCCACACCCAGGGCAAGTGGAGAGGGAGCTG 423  
QY 121 GlyThrAspLeuValSerIleProHisGlyProAsnValThrValArgAlaAsnIleAla 140  
Db 424 GGACCCAGCTGTGTACATCTCCCATGGCCCCCAACCTCATGCTGTGCAACATTTGCT 483  
QY 141 AlaIleThrGlnSerAspLysPhePheIleAsnGlySerAsnTyrPGLuGlyIleLeuGly 160  
Db 484 GCCATCATGTGAATGACACAAAGTTCTTCAACAGGCTCCAACTGGAGAGGACATCTGGGG 543  
QY 161 LeuAlaTyrAlaGluIleAlaArgProAspAspSerLeuGluProPhePheAspSerLeu 180  
Db 544 CTGGCTTATGCTGAGATTGGCCAGGCTGACGACTCCCTGGAGCTTTCTTGACTCTCTG 603  
QY 181 ValIleGlnThrHisValProAsnLeuPheSerLeuGlnLeuGlyValAlaGlyPhePro 200  
Db 604 GTAAAGACAGACCCAGTCCCAACCTCTCTCCCTGAGCTTGTGTGTGCTGTGCTTCC 663  
QY 201 LeuAsnGlnSerGluValLeuAlaSerValGlyGlySerMetIleIleGlyGlyIleAsp 220  
Db 664 CTCACACAGCTGTAAGTGTGGCTGTGTGAGAGGAGATGATCATTTGAGATATCGAC 723  
QY 221 HisSerLeuTyrThrGlySerLeuTyrTyrThrProIleArgArgGluTyrTyrGlu 240  
Db 724 CACTCGGTATCACAGGCACTCTGTGTATCACCATCCGGGGAGGTGTATTTATGAG 783  
QY 241 ValIleIleValArgValGluIleAsnGlyGlnAspLeuLysMetAspCysLysGluTyr 260  
Db 784 GTCATCATTTGTGGGGTGGAGATCATGACAGAGATCTGAAAATGAGACTGCAAGAGTAC 843  
QY 261 AsnTyrAspLysSerIleValAspSerGlyTyrThrAsnLeuArgLeuProIlyLysVal 280  
Db 844 AACCTATGACAAAGCATTTGTGACAGTGGCACCAACCACTTGTGGCCAAAGAAAGT 903  
QY 281 PheGluAlaAlaValLysSerIleLysAlaIleAsnSerThrGlnLysPheProAspGly 300  
Db 904 TTTGAAAGCTGCACTCAATTCATCAAGGACGCTCTCTCCAGAGAAAGTTCCCTATGAT 963  
QY 301 PheTyrLeuGlyGluGlnLeuValCysTyrPGLuAlaGlyThrThrProTyrAsnIlePhe 320  
Db 964 TTCTGGCTAGAGAGAGCTGGTGTGTCTGGCAAGACAGCACCCCTTGGAACTTTTC 1023  
QY 321 ProValIleSerLeuTyrIleuMetGlyGluValThrAsnGlnSerPheArgIleThrIle 340  
Db 1024 CCAATCATCTCACTTAATGAGGTGAGTTTACCAACAGTCTTCGCGCATCACCATC 1083  
QY 341 LeuProGlnGlnTyrLeuArgProValGluAspValAlaThrSerGlnAspAspCysTyr 360  
Db 1084 CTTCCGAGACATACCTCGGCGCAATGAGATGTGGCCACTCCCAAGACACTGTTTAC 1143  
QY 361 LysPheAlaIleSerGlnSerSerThrGlyThrValMetGlyAlaValIleuMetGlnGly 380  
Db 1144 AAGTTTGCATCTCAAGTCATCCAGGCGCACTGTTATGAGAGCTGTTATCATGAGAGGC 1203  
QY 381 PheTyrValValPheAspArgAlaArgLysArgIleGlyPheAlaValSerIleCysHis 400  
Db 1204 TTCTACGTTTGTGTGATCGGGCCGAAACCAATTTGCTTGTGTAGAGGCTTGCAT 1263  
QY 401 ValHisAspGluPheArgThrAlaAlaValGluGlyProPheValThrLeuAspMetGlu 420  
Db 1264 GTGACAGATGATTCAGGACGCGACGCGTGGAGAGGCTTTTGTGCACTTGGACATGGA 1323  
QY 421 AspCysGlyTyrAsnIleProGlnThrAspGluSerThr 433  
Db 1324 GACTGTGCTACAACTTCCACAGACAGATGATGATCAC 1362  
RESULT 11  
US-09-548-366F-3

/ Sequence 3, Application US/09548366F  
/ Patent No. 6797487  
/ GENERAL INFORMATION:  
/ APPLICANT: GURNEY ET AL.  
/ TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES  
/ FILE REFERENCE: 29915/6280F  
/ CURRENT APPLICATION NUMBER: US/09/548,366F  
/ PRIOR FILING DATE: 2000-04-12  
/ PRIOR APPLICATION NUMBER: US 60/155,493  
/ PRIOR FILING DATE: 1999-09-23  
/ PRIOR APPLICATION NUMBER: US 09/404,133  
/ PRIOR FILING DATE: 1999-09-23  
/ PRIOR APPLICATION NUMBER: PCT/US99/20881  
/ PRIOR FILING DATE: 1999-09-23  
/ PRIOR APPLICATION NUMBER: US 60/101,594  
/ PRIOR FILING DATE: 1998-09-24  
/ NUMBER OF SEQ ID NOS: 73  
/ SOFTWARE: Patentin version 3.1  
/ SEQ ID NO 3  
/ LENGTH: 2070  
/ TYPE: DNA  
/ ORGANISM: Homo sapiens  
US-09-548-366F-3  
Alignment Scores:  
Pred. No.: 4,34e-257 Length: 2070  
Score: 2267.00 Matches: 429  
Percent Similarity: 99.31% Conservative: 1  
Best Local Similarity: 99.08% Mismatches: 3  
Query Match: 99.08% Indels: 0  
DB: 4 Gaps: 0  
US-10-726-967a-84 (1-433) x US-09-548-366F-3 (1-2070)  
QY 1 ThrGlnHisGlyIleArgLeuProLeuArgSerGlyLeuGlyGlyAlaProLeuGlyLeu 20  
Db 64 ACCAGACAGGACATCCGGCTGCTCCCTGCGACGGGCTGGGGGGCGCCCCCTGGGGCTG 123  
QY 21 GluIleAsnLeuGluThrAspGluGluProGluGluProGlyLysArgArgLysPheVal 40  
Db 124 CGGCTGCCCCGAGAGACCGACGAAGAGCCGAGAGAGCCGGCGGAGGGGACCTTTGTG 183  
QY 41 GluMetValAspAsnLeuArgGlyLysSerGlyGlnGlyTyrTyrValGluMetThrVal 60  
Db 184 GAGATGATGACAACTGAGGGGCAAGTGGGGCAGGGCTACTACGTGAGATGACCGTG 243  
QY 61 GlySerProProGlnThrLeuAsnIleLeuValAspThrGlySerSerAsnPheAlaVal 80  
Db 244 GCGAGCCCCCGAGAGCGTCAACATCTGTGTGATACAGGACGACGTAACTTGCAGTG 303  
QY 81 GlyAlaAlaProHisProPheLeuHisArgTyrTyrGlnArgGlnLeuSerSerThrTyr 100  
Db 304 GGTGTGCCCCCACCCTTCTGCACTGCTACTACCAAGGACAGCTGTCCAGACATAC 363  
QY 101 ArgAspLeuArglysglyValTyrValProTyrThrGlnGlyLysTyrPGLuGlyGluLeu 120  
Db 364 CGGAGACCTCCGGAGGGGTGTATGTGCTCCACACCCAGGGCAAGTGGAGAGGGAGCTG 423  
QY 121 GlyThrAspLeuValSerIleProHisGlyProAsnValThrValArgAlaAsnIleAla 140  
Db 424 GGACCCAGCTGTGTACATCTCCCATGGCCCCCAACCTCATGCTGTGCAACATTTGCT 483  
QY 141 AlaIleThrGlnSerAspLysPhePheIleAsnGlySerAsnTyrPGLuGlyIleLeuGly 160  
Db 484 GCCATCATGTGAATGACACAAAGTTCTTCAACAGGCTCCAACTGGAGAGGACATCTGGGG 543  
QY 161 LeuAlaTyrAlaGluIleAlaArgProAspAspSerLeuGluProPhePheAspSerLeu 180  
Db 544 CTGGCTTATGCTGAGATTGGCCAGGCTGACGACTCCCTGGAGCTTTCTTGACTCTCTG 603  
QY 181 ValIleGlnThrHisValProAsnLeuPheSerLeuGlnLeuGlyValAlaGlyPhePro 200

```

Db      604 GTAAGCAGACCCAGCTTCCCAACTCTTCTCCGACGCTTGTGCTGCTGCTTCCCC 663
Qy      201 LeuaengInserGluValleuaIaSerValGlyGlySerMetIleIleGlyGlyIleap 220
Db      664 CTCACACAGACTGAGAGTCTGGCTCTCTGCGAGGGAGCATATCATATGAGATATCGAC 723
Qy      221 HisSerLeuTyThrGlySerLeuTpyTyThrProIleArgArgIleuTpyTyTyrglu 240
Db      724 CACTGCTGTACACAGCAGCTCTGTGTATACCCATCCGCGGGAGGTATATATGAG 783
Qy      241 ValIleIleValArgValGluIleAengIyGlnAapLeuIleuMetAapCySlySgluTy 260
Db      784 GTTCATCATTTGTGGGTGAGATCATATGACAGATCTGAAAATGACATGCAAGAGATAC 843
Qy      261 AenTyAspLySerIleValAapSerGlyThrThrAsnLeuArgLeuProIlySlyVal 280
Db      844 AACTATGACAGAGCATTTGTGACAGTGGCACACCACTTCGTTTGGCCCAAGAAAGTG 903
Qy      281 PheGluAlaIleValIySerIleIyValAaIaSerSerThrGluSlyPheProAapGly 300
Db      904 TTTGAAGCTGACGTCAATTCATCAAGCAGCCTCTCCACGAGAAAGTTCCCTGATGGT 963
Qy      301 PheTyrLeuGlyGluGlnIleuValCysTTPGlnAaGlyThrThrProTyrAsnIlePhe 320
Db      964 TTCTGCTGAGAGAGAGCTGTGTCTGCGACAGCAGCCCTTGGAACATTTTC 1023
Qy      321 ProValIleSerLeuTyThrLeuMetGlyGluValThrAsnGlnSerPheArgIleThrIle 340
Db      1024 CCAATCATCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTC 1083
Qy      341 LeuProGlnGlnTyThrLeuArgProValGluAapValAaIaThrSerGlnAapPcySlyTy 360
Db      1084 CTTCCTCCACCAATACCTCGCGCCAGAGAGATGTGGCCACGTCACCAAGACGACTGTTCAC 1143
Qy      361 LysPheAlaIleSerGlnSerSerThrGlyThrValMetGlyValIleMetGlyGly 380
Db      1144 AAGTTGCCATCTCACAGTATCATCAGGGCATCTGTTATGAGACTTTATCATGAGGGGC 1203
Qy      381 PheTyValIleValPheAspArgAlaArgIySlyArgIleGlyPheAlaValSerAlaCySly 400
Db      1204 TTCTACGTTGTCTTTATCGGGCCGAAAACGAATTTGGCTTGTCTGTGACGCTTGCAT 1263
Qy      401 ValHisAapGluPheArgThrAlaIleValGlyGlyProPheValThrLeuAapMetGly 420
Db      1264 GTGCACGATGAGTTCAGGACGCGCAGCGGTGGAAGGCCCTTTGTCACTTGGACATGGA 1323
Qy      421 AapCyGlyTyThrAsnIleProGlnThrAapGlnSerThr 433
Db      1324 GACTGTGCTACAACTTCACAGACGATGATCAACC 1362

RESULT 12
US-09-548-368D-3
; Sequence 3, Application US/09548368D
; Patent No. 6825023
; GENERAL INFORMATION:
; APPLICANT: GUNNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 29915/6280C
; CURRENT APPLICATION NUMBER: US/09/548,368D
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 2070

```

```

; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-548-368D-3
Alignment Scores:
Pred. No.: 4,34e-257 Length: 2070
Score: 2267.00 Matches: 429
Percent Similarity: 99.31% Conservative: 1
Best Local Similarity: 99.08% Mismatches: 3
Query Match: 99.08% Indels: 0
DB: 4 Gaps: 0
US-10-726-967a-84 (1-433) x US-09-548-368D-3 (1-2070)
Qy      1 ThrGlnHisGlyTylleArgLeuProleuArgSerGlyLeuGlyValAProleuGlyLeu 20
Db      64 ACCCAGCAGCGGATCCGCTGCTGCGACAGGCGCTGGGGGGGCCCCCTGGGGGCTG 123
Qy      21 GlnIleAenLeuGluThrAapGluGluProGluGluProGlyArgArgGlySerPheVal 40
Db      124 CCGCTGCCCCGGAGACCGACAGAAAGCCGAGAGCCCGCGCGAGGGGAGGCTTTGG 183
Qy      41 GluMetValAapAsnLeuArgGlyIySserGlyGlnGlyTyTyTyValGluMetThrVal 60
Db      184 GAGATGTGGAACAACCTGAGGGGCAAGTCCGGGCGAGGCTACTTACGTGGAGATGACCGTG 243
Qy      61 GlySerProProGlnThrLeuAsnIleLeuValAapThrGlySerSerAspPheAlaVal 80
Db      244 GGCACCCCCCCACACAGCTCAACATCTGTGTGATACAGGAGAGAGATGATCTTGCAGTG 303
Qy      81 GlyAlaIleProHisProPheLeuHisArgTyTyTyGlnArgGlnLeuSerSerThrTy 100
Db      304 GGTGTGCCCCCACCACCCCTTCTGTGATGCTACTACAGAGGCACTGTCCAGCATATAC 363
Qy      101 ArgAapLeuArgIyGlyValTyValProTyTyThrGlnGlyIySerTTPGlyGlyLeu 120
Db      364 CCGGACCTCCGAAAGGTGTGATGTGCTTACACCCAGGCGAAGTGGGAAGGGAGCTG 423
Qy      121 GlyThrAapLeuValSerIleProHisGlyProAenValThrValArgAlaAsnIleAla 140
Db      424 GGCACGACCTGTGAAGCATCCCAATGCCCCCAAGCTCACTGTGTGTCACCAATGTCT 483
Qy      141 AlaIleThrGlnSerAspPhePheIleAengIySerAsnTTPGlyIleLeuGly 160
Db      484 GGCATACATGAATCAGACAAATTTCTTATCAACGGCTCCAACTGGGAAGGATCTCGGG 543
Qy      161 LeuAlaTyAlaGluIleAlaArgProAapSerSerLeuGluProPhePheAapSerLeu 180
Db      544 CTGGCCTATGTGAGATTTGCCAGGCTGACACTCCCTGGAGCTTTCTTGACTCTG 603
Qy      181 ValIyGlnThrHisValProAenLeuPheSerLeuGlnLeuCySlyAlaGlyPhePro 200
Db      604 GTAAGCAGACCCAGCTTCCCAACTCTTCTCCGACGCTTGTGCTGCTGCTTCCCC 663
Qy      201 LeuaengInserGluValleuaIaSerValGlyGlySerMetIleIleGlyGlyIleap 220
Db      664 CTCACACAGACTGAGAGTCTGGCTCTCTGCGAGGGAGCATATCATATGAGATATCGAC 723
Qy      221 HisSerLeuTyThrGlySerLeuTpyTyThrProIleArgArgIleuTpyTyTyrglu 240
Db      724 CACTGCTGTACACAGCAGCTCTGTGTATACCCATCCGCGGGAGGTATATATGAG 783
Qy      241 ValIleIleValArgValGluIleAengIyGlnAapLeuIleuMetAapCySlySgluTy 260
Db      784 GTTCATCATTTGTGGGTGAGATCATATGACAGATCTGAAAATGACATGCAAGAGATAC 843
Qy      261 AenTyAspLySerIleValAapSerGlyThrThrAsnLeuArgLeuProIlySlyVal 280
Db      844 AACTATGACAGAGCATTTGTGACAGTGGCACACCACTTCGTTTGGCCCAAGAAAGTG 903
Qy      281 PheGluAlaIleValIySerIleIyValAaIaSerSerThrGluSlyPheProAapGly 300
Db      904 TTTGAAGCTGACGTCAATTCATCAAGCAGCCTCTCCACGAGAAAGTTCCCTGATGGT 963

```



QY 301 PheTrpLeuGlyGluGlnLeuValCysTrpGlnAlaGlyThrThrProTrpAsnIlePhe 320  
Db 964 TTCTGGCTAGAGAGAGAGCTGTGTGGCGACAGGACACACCCCTGGAACTTTTC 1023  
QY 321 ProValIleSerLeuTyTrLeuMetGlyGluValThrAsnGlnSerPheArgIleThrIle 340  
Db 1024 CCAAGTCATCTCACTCACTAAATGGGTGAGTTTCCAAACCAAGTCCTTCGCCATCACATC 1083  
QY 341 LeuProGlnGlnTyTrLeuArgProValGluAspValAlaThrSerGlnAspAspCysTyTr 360  
Db 1084 CTTCGGACGAAATACCTGGCGGCGAGTGAATGTGGCCACGTCCCAAGACGACTGTAC 1143  
QY 361 LysPheAlaIleSerGlnSerSerThrGlyThrValMetGlyAlaValIleMetGluGly 380  
Db 1144 AAGTTTGCCATCTCAACAGTCATCCACGGGCACTTTTAGAGCTGTATCATGAGGGC 1203  
QY 381 PheTyValValPheAspArgAlaArgLysArgIleGlyPheAlaValSerAlaCysHis 400  
Db 1204 TTCTACCTGTCTTGTGATCGGGCCGAAAGCAATTGGCTTCTGTCAAGCGCTTCAN 1263  
QY 401 ValHisAspGluPheArgThrAlaAlaValGluGlyProPheValThrLeuAspMetGlu 420  
Db 1264 GTGACGATAGATTCAAGAGGCGAGCGGTGAAGGCCCTTTGTCACTTGACATGGAA 1323  
QY 421 AspCysGlyTyTrAsnIleProGlnThrAspGluSerThr 433  
Db 1324 GACTGTGGCTAACAACTTCACAGACGATGACTCAAC 1362

RESULT 13  
US-09-794-925A-3  
/ Sequence 3, Application US/09794925A  
/ Patent No. 6828117  
/ GENERAL INFORMATION:  
/ APPLICANT: Gurney et al.  
/ TITLE OF INVENTION: Alzheimer's Disease Secretase, APP Substrates Therefor and Uses  
/ FILE REFERENCE: 29915/6280H1  
/ CURRENT FILING DATE: 2001-02-27  
/ PRIOR APPLICATION NUMBER: 09/416,901  
/ PRIOR FILING DATE: 1999-10-13  
/ PRIOR APPLICATION NUMBER: 60/155,493  
/ PRIOR FILING DATE: 1999-09-23  
/ PRIOR APPLICATION NUMBER: 09/404,133  
/ PRIOR FILING DATE: 1999-09-23  
/ PRIOR APPLICATION NUMBER: PCT/US99/20881  
/ PRIOR FILING DATE: 1999-09-23  
/ PRIOR APPLICATION NUMBER: 60/101,594  
/ NUMBER OF SEQ ID NOS: 74  
/ SOFTWARE: PatentIn Ver. 2.0  
/ SEQ ID NO 3  
/ LENGTH: 2070  
/ TYPE: DNA  
/ ORGANISM: Homo sapiens  
US-09-794-925A-3

Alignment Scores:  
Pred. No.: 4,34e-257 Length: 2070  
Score: 2267.00 Matches: 429  
Percent Similarity: 99.31% Conservative: 1  
Best Local Similarity: 99.08% Mismatches: 3  
Query Match: 99.08% Indels: 0  
DB: 4 Gaps: 0

US-10-726-967A-84 (1-433) x US-09-794-925A-3 (1-2070)

QY 1 ThrGlnHisGlyTAlaArgLeuProLeuArgSerGlyLeuGlyAlaProLeuGlyLeu 20  
Db 64 ACCCGACAGCGGATCCGGCTGCCCTGGCGACGGGCTGGGGGGGCCCCCTGGGGGCTG 123  
QY 21 GluIleAsnLeuGluThrAspGluGluProGluGluProGluGlyArgAlaCysHis 40

Db 124 CGGCTGCCCCGGGAAACCGGACGAAAGGCCGAGGAGCCCGGCGGAGGGGACGCTTGTG 183  
QY 41 GluMetValAspAsnLeuArgGlyLysSerGlyGlnGlyTyTrValGluMetThrVal 60  
Db 184 GAGAGTGTGACAACTTAAAGGGGCAAGTCCGGGGCAAGGCTTCACTGAGATACCGTGG 243  
QY 61 GlySerProProGlnThrLeuAsnIleLeuValAspThrGlySerSerAspAsnAlaVal 80  
Db 244 GGCAGCCCCCGGACGAGACGCTCAACATCTGTGTGATACAGGACAGACAGTAATTGGAGTG 303  
QY 81 GlyAlaAlaProHisAspPheLeuHisArgGlyTyTrGlnArgGlnLeuSerSerThrTyTr 100  
Db 304 GGTGTGCCCCCGACCCCTTCTGTCACTCACTTACCAAGGCGAGCTGTCCAGCAATAC 363  
QY 101 ArgAspLeuArgLysGlyValTyTrValProTyThrGlnGlyLysTrpGluGlyLeu 120  
Db 364 CGGAGCTCCGGAAAGGGTGTATAGTCCCTTACACCAAGGCGAAGTGGAAAGGAGACTG 423  
QY 121 GlyThrAspLeuValSerIleProHisGlyProAsnValThrValArgAlaAsnIleAla 140  
Db 424 GGCACCGGACTGTGTAGGATCCCAATGGCCCAAGCTCACTGTGCGTCCAAATTCCT 483  
QY 141 AlaIleThrGlnSerAspLysPhePheIleAsnGlySerSerThrProGluGlyIleLeuGly 160  
Db 484 GCCATCACTGAATACAGACAGTCTTATCATCAAGGCTTCACTGGAGAGGCACTTGGGG 543  
QY 161 LeuAlaTyAlaGluIleAlaArgProAspAspSerLeuGluProPheAspSerLeu 180  
Db 544 CTGGCTATGTGTGATTTGCCAGAGCTGACGATCTCCCTGGAAGCTTCTTGTACTCTGTG 603  
QY 181 ValLysGlnThrHisValProAsnLeuPheSerLeuGlnLeuGlyGlyAlaGlyPhePro 200  
Db 604 GTAAGGAGAGCCAGTTCACCTCTTCTCCAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGT 663  
QY 201 LeuAsnGlnSerGluValLeuAlaSerValGlySerMetIleIleGlyGlyIleAsp 220  
Db 664 CTCAACAGAGCTGAAGT 723  
QY 221 HisSerLeuTyTrThrGlySerLeuTyTrTyTrProIleArgArgGluTyTrTyTrGlu 240  
Db 724 CACTGCTGTACACAGGAGTCTCTGTATACACCATCCGGGGGAGTGTATATGAG 783  
QY 241 ValIleIleValArgValGluIleAsnGlyGlnAspLeuLysMetAspCysLysGlyTyTr 260  
Db 784 GTCACTATGTGGGGGTGAGATCAATGACAGGATCTAAATGATGACGCAAGAGTAC 843  
QY 261 AsnTyTrAspLysSerIleValAspSerGlyThrThrAsnLeuArgLeuProLysVal 280  
Db 844 AACTATGACAAAGAGCATTTGTGACATGTGACACCAACCTTCCTTGTGCCAAGAAAGTG 903  
QY 281 PheGluAlaAlaValLysSerIleLysAlaAlaSerSerThrGluLysPheProAspGly 300  
Db 904 TTGTAAGCTGCACTCAATCAATCAAGGCGCTCTCTCCACGGAAGTAAGTTCCCTGATGAT 963  
QY 301 PheTrpLeuGlyGluGlnLeuValCysTrpGlnAlaGlyThrThrProTrpAsnIlePhe 320  
Db 964 TTCTGGCTAGAGAGAGAGCTGTGTGGCGACAGGACACACCCCTGGAACTTTTC 1023  
QY 321 ProValIleSerLeuTyTrLeuMetGlyGluValThrAsnGlnSerPheArgIleThrIle 340  
Db 1024 CCAAGTCATCTCACTCACTAAATGGGTGAGTTTCCAAACCAAGTCCTTCGCCATCACATC 1083  
QY 341 LeuProGlnGlnTyTrLeuArgProValGluAspValAlaThrSerGlnAspAspCysTyTr 360  
Db 1084 CTTCGGACGAAATACCTGGCGGCGAGTGAATGTGGCCACGTCCCAAGACGACTGTAC 1143  
QY 361 LysPheAlaIleSerGlnSerSerThrGlyThrValMetGlyAlaValIleMetGluGly 380  
Db 1144 AAGTTTGCCATCTCAACAGTCATCCACGGGCACTTTTAGAGCTGTATCATGAGGGC 1203  
QY 381 PheTyValValPheAspArgAlaArgLysArgIleGlyPheAlaValSerAlaCysHis 400



Db 1204 TTCTACGTTGCTTGTGATCGGGCCGGAACCAATTGGCTTGTCTGACGCGCTTGCCAT 1263  
 Qy 401 VALHISAPGLUPHeargThra1a1aValGluGlyProPheValThrLeuaspMetGlu 420  
 Db 1264 GTGCACGATGATTGACGAGCGGAGGCGGTGGAAGGCCCTTTGTGACACTTGGACATGGAA 1323  
 Qy 421 AepCyeglyTYrAsn1LeProGlnThraSpG1uSerThr 433  
 Db 1324 GACTGTGGCTACAACTTCACAGACAGATGAGTCAACC 1362

RESULT 14

US-09-806-194A-3  
 / Sequence 3, Application US/09806194A  
 / Patent No. 6835565  
 / GENERAL INFORMATION:  
 / APPLICANT: Gurney, Mark E.  
 / APPLICANT: Bienkowski, Michael J.  
 / APPLICANT: Heinrichson, Robert L.  
 / APPLICANT: Parodi, Luis A.  
 / APPLICANT: Yan, Riqiang  
 / APPLICANT: Pharmacia & Upjohn Company  
 / TITLE OF INVENTION: Alzheimer's Disease Secretase  
 / FILE REFERENCE: 6177.P CP  
 / CURRENT APPLICATION NUMBER: US/09/806,194A  
 / CURRENT FILING DATE: 2001-09-17  
 / PRIOR APPLICATION NUMBER: 60/101,594  
 / PRIOR FILING DATE: 1998-09-24  
 / NUMBER OF SEQ ID NOS: 49  
 / SOFTWARE: PatentIn Ver. 2.0  
 / SEQ ID NO 3  
 / LENGTH: 2070  
 / TYPE: DNA  
 / ORGANISM: Homo sapiens  
 US-09-806-194A-3

Alignment Scores:  
 Pred. No.: 4,346-257 Length: 2070  
 Score: 2267.00 Matches: 429  
 Percent Similarity: 99.31% Conservative: 1  
 Best Local Similarity: 99.08% Mismatches: 3  
 Query Match: 99.08% Indels: 0  
 DB: 4 Gaps: 0

US-10-726-967A-84 (1-433) x US-09-806-194A-3 (1-2070)

Qy 1 ThhGlnHISGLYIleargLeuPProLeuArgSerGlyLeuGlyValAlaProLeuGlyLeu 20  
 Db 64 ACCGACGACGGCAATCCGGCTGCGGACGCGGCTGGGGGGCCGCCCCCTGGGGCTG 123  
 Qy 21 GlnIleAsnLeuGluThraSpG1uGluProGluGluProGlyValArgArgGlySerPheVal 40  
 Db 124 CGGCTGCCCCGGAGACCGACGAAGAGCCCGAGAGCCGCGGCGGACGCTTTGTC 183  
 Qy 41 GlnMetValAspAsnLeuArgGlyLeuSerGlyGlnGlyTyrTyrValGlnMetThrVal 60  
 Db 184 GAGATGGTGACAACTCGAGGGCAAGTCCGGGCGAGGGCTACTACGAGATGACCGTG 243  
 Qy 61 GlySerProProGlnThraLeuAsn1LeuValAspThrGlySerSerAsnPheAlaVal 80  
 Db 244 GGGAGCCCCCGGACGAGCTCAATCTGATGATACAGGACGACGACTTACTTTGACGTG 303  
 Qy 81 GlyAlaAlaProHisProPheLeuHisArgTyrTyrGlnArgGlnLeuSerSerThrTyr 100  
 Db 304 GGTGTGTCGCCGCCACCCCTTCCTGACATCGCTACTACAGAGGCGGTGTCACAGACATAC 363  
 Qy 101 ArgAspLeuArgGlyValTyrValProTyrThrGlnGlyTyrTyrGlnGlyGluLeu 120  
 Db 364 CGGGAAGCTCCGGAAGGGGTGTGTATGTCCCTTACACCCAGGGCAAGTGGGAAGGGAGCTG 423  
 Qy 121 GlyThraSpLeuValSer1LeuProHisArgGlyProAsnValThrValArgAlaAsn1Lea 140  
 Db 424 GGCACCGAAGCTGGTAAAGATCCCGCATGGCCCCCAAGCTCACTGTGCTGCCAATGCT 483

Qy 141 AlaIleThrGluSerAspPhePhe1IleAsnGlySerAsnTPrgIuGlyIleLeuGly 160  
 Db 484 GCCATCACTGATACAGACAAATTCTTATCAACGGCTCAACTGGAAAGCATCTGGGG 543  
 Qy 161 LeuAlaTyrAlaGluIleAlaArgProAspAspSerLeuGluProPhePheAspSerLeu 180  
 Db 544 CTGGCTATGCTGAGATTGCCAGGCGCTGACCACTCCCTGGAGCCCTTCTTGACTCTCTG 603  
 Qy 181 ValIleGlnThrHisValProAsnLeuPheSerLeuGlnLeuCySGlyAlaGlyPhePro 200  
 Db 604 GTAAAGCAGACCACGGTCCCAACTCTTCTCCGACGCTTGTGGTGGCTTCTCC 663  
 Qy 201 LeuAsnGlnSerGluValIleuAlaSerValGlyGlySerMetIleIleGlyIleAsp 220  
 Db 664 CTCACCGAGTGTGAAGTCTGGCTCTGTCCGAGGAGACATGATATTGAGATTCAC 723  
 Qy 221 HisSerLeuTyrThrGlySerLeuTPyrThrProIleArgArgGluTPyrTyrGlu 240  
 Db 724 CACTGCTGTACACAGGAGTCTGTGATACACCATCCGCGGAGGTGATATTATAG 783  
 Qy 241 ValIleIleValArgValGluIleAsnGlyGlnAspLeuIleuAspCySGlyAspGlyTyr 260  
 Db 784 GTCATCATTTGTCCGGGTGAGATCATATGACAGATCTGAANAATGACTGCAAGAGTAC 843  
 Qy 261 AsnTyrAspIleSerIleValAspSerGlyThrThrAsnLeuArgLeuProIleVal 280  
 Db 844 AACATGACAAAGCATTTGTGACGTGGACACCAACCTTCCTTGGCCCAAGAAAGT 903  
 Qy 281 PheGluAlaAlaValIleSerIleValAspSerGlyThrThrAsnLeuArgLeuProIleVal 300  
 Db 904 TTTGAAGCTGACGTCATATTCATCAAGGACGCTCTCCACGAGAAAGTCCCTATGT 963  
 Qy 301 PheTPLeuGlyGluGlnLeuValCySerTPGlnAlaGlyThrThrProThrAsn1IlePhe 320  
 Db 964 TTCTGGCTAGGACACAGCTGTGTGTGCTGGCAAGGACCAACCCCTTGAACATTTTC 1023  
 Qy 321 ProValIleSerLeuTyrIleuMetGlyGluValThrAsnGlnSerPheArgIleThrIle 340  
 Db 1024 CCAAGTCATTCCTCTTACTTAATGGGTGAGTTACCAACCACTCTTCGCAATCACAC 1083  
 Qy 341 LeuProGlnGlnTyrIleuArgProValGluAspValAlaThrSerGlnAspAspCySlyr 360  
 Db 1084 CTTCCGACGACATACCTGGGCGCAGTGGAAAGTGTGCCACGCTCCAGACACATGTTAC 1143  
 Qy 361 LysPheAlaIleSerGlnSerSerThrGlyThrValMetGlyAlaValIleMetGluGly 380  
 Db 1144 AAGTTGCACTCTCAAGTCATCCAGCGGACCTGTATGGAGCTGTATCATGGAGGCG 1203  
 Qy 381 PheTyrValValPheAspArgAlaArgLysArgIleGlyPheAlaValSerAlaCyHis 400  
 Db 1204 TTCTACGTTGCTTGTGATCGGCGCGAACAATGGCTTTGCTGTCAAGCGCTTGCCAT 1263  
 Qy 401 ValHISAPGLUPHeargThra1a1aValGluGlyProPheValThrLeuaspMetGlu 420  
 Db 1264 GTGCACGATGATTGACGAGCGGAGGCGGTGGAAGGCCCTTTGTGACACTTGGACATGGAA 1323  
 Qy 421 AepCyeglyTYrAsn1LeProGlnThraSpG1uSerThr 433  
 Db 1324 GACTGTGGCTACAACTTCACAGACAGATGAGTCAACC 1362

RESULT 15

US-09-724-566A-42  
 / Sequence 42, Application US/09724566A  
 / Patent No. 6627739  
 / GENERAL INFORMATION:  
 / APPLICANT: Anderson, John P.  
 / APPLICANT: Basi, Gurigbal  
 / APPLICANT: Doane, Minh Tam  
 / APPLICANT: Frigon, No. 6627739mand  
 / APPLICANT: John, Varghese  
 / APPLICANT: Power, Michael  
 / APPLICANT: Sinha, Sukanto  
 / APPLICANT: Tatsuno, Gwen

```
/ APPLICANT: Tung, Jay
/ APPLICANT: Mang, Shuwen
/ TITLE OF INVENTION: Beta-Secretase Enzyme Compositions and
/ FILE REFERENCE: 228-US-NEWC2
/ CURRENT FILING DATE: 2000-11-28
/ PRIOR APPLICATION NUMBER: US 09/724,566A
/ PRIOR FILING DATE: 2000-02-10
/ PRIOR APPLICATION NUMBER: 60/119,571
/ PRIOR FILING DATE: 1999-02-10
/ PRIOR APPLICATION NUMBER: 60/139,172
/ PRIOR FILING DATE: 1999-06-15
/ NUMBER OF SEQ ID NOS: 104
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO: 42
/ LENGTH: 2348
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-724-566A-42

Alignment Scores:
Pred. No.: 5,33e-257 Length: 2348
Score: 2267.00 Matches: 429
Percent Similarity: 99.31% Conservative: 1
Best Local Similarity: 99.08% Mismatches: 3
Query Match: 99.08% Indels: 0
DB: 4 Gaps: 0

US-10-726-967a-84 (1-433) x US-09-724-566A-42 (1-2348)

QY 1 THRGINHLEGLYTLAARGLEUProLeuWARGSERGLYLEGLYVALAProLeuGLYLEU 20
DB 369 ACCCAGCAGGAGCATCGGCTGCCCTGCGCAGCGGCTGAGGCGCCCTTGAGGCTG 428
QY 21 GIUILEuLeuGLIThraSPGLUGIUPProGLUGIUPProGLIYArgArgLYSerPheVal 40
DB 429 CGGCTGCCCGGGAAGACCAAGACGCGGAGACCGGCGGAGGGGCGAGCTTTGTG 488
QY 41 GIUetValASPAsnLeuArgLYLysSerGLYGLYTYRValGLUetThrVal 60
DB 489 GAGATGATGACAACTGAGGGGCAAGTCGGGGGAGGCTACTAGTGAGATGACCGTG 548
QY 61 GLYSerProProGLIInThrieuSnlleuValAAPTThrGLYSerSerAsnPhelVal 80
DB 549 GGCACCGCCCGCAGACGCTCAACATCCGTGATGATACAGCAGCAGATTAACCTTTCAGTG 608
QY 81 GIYAlAAlProHISProPheLeuHISArgTYRGLYArgGLInLeuSerSerThrTYR 100
DB 609 GGTGCTGCCCGCCACCCCTTCTGCACTGCTACACAGAGGCACTGTCCAGCACAATAC 668
QY 101 ArgASPLeuArgLYGLYValTYRValProLYRThrGLYLYSTPGLUGIULEU 120
DB 669 CGGAGACCTCCGGAAAGGTGTATGTGCCCTACACCCAGGCAAGTGGGAAGGGAGCTG 728
QY 121 GLYThraSPLeuValSerTleProHISGLYProAsnValThrValArgAlaAsnTleAla 140
DB 729 GGCACCGACCTGTGTAGACATCCCGCAAGCTCACTGTCGTGCCAACAATGTCT 788
QY 141 AlaIleThrGLYSerASPLeuPhePheIleAsnGLYSerAsnTProGLUGIYIleuGLY 160
DB 789 GCCATCACTGAATCAAGCAAGTTCTTCATCAACGCGCTCCACTGGGAAGGCACTCTGGGG 848
QY 161 LeuAlaTYRAlaGLIleAlaArgProASPAspSerLeuGLUProPhePheASPserLeu 180
DB 849 CTGGCCTATGTGATGATGCGAGGCTGACGACTCCCTGAGAGCTTTCTTGACCTCTCG 908
QY 181 ValLYSGIInThrHISValProAsnLeuPheSerLeuGLInLeuCYGLYAlaGLYphePro 200
DB 909 GTTAAAGCAGACCAAGTTCCTTCTTCTCCCTGACAGCTTTGTGTGCTGGCTTCTCC 968
QY 201 LeuAsnGLInSerGLIValLeuAlaSerValGLYGLYSerMetIleGLYGLYIleASP 220
```

```
DB 969 CTCAACGAGCTGAAGTCTGGCTCTGTGCGAAGGAGCATGATCATTTGAGGATGAC 1028
QY 221 HISerLeuTYRThrGLYSerLeuTPTYRThrProIleArgArgLUPTTYRGLU 240
DB 1029 CACTGCTGTACACAGGAGCTCTGGTATACACCAATCCGGCGGAGTGATTAATGAG 1088
QY 241 ValIleIleValAlaValAlaGLIleAsnGLYInASPLeuYMetASPValYGLUTyr 260
DB 1089 GTGATCATTTGCGAGTGAAGATCAATGACAGGATCTGAAAATGACTGCAGAGATAC 1148
QY 261 AsnTYRASPLeuSerIleValASPserGLYThrThraSPLeuArgLeuProLYLeVal 280
DB 1149 AACATGACAGAGACATTTGTGACAGTGGCACCAACCACTTGTCTTGCCCAAGAAATG 1208
QY 281 PheGLUAlaAlaValLYSerIleLYAlaIleAserThrGLUYPheProASPGLY 300
DB 1209 TTGAGAGCTGACGACAAATCCATCAAGCAGCCCTCCACGAGAAATTCCTGATGATGT 1268
QY 301 PheTPLeuGLYGLUGIInLeuValCYETProGLInAlaLYThrThrProTPAsnTlePhe 320
DB 1269 TTCTGGCTAGGAGAGAGCTGTGTGCTGGCAGACAGCAGCACCCCTTGAGACATTTTC 1328
QY 321 ProValIleSerLeuTYRLeuMetGLYValThraSPInserPheArgIleThrIle 340
DB 1329 CCACTCATCTCATCTACCTTAATGGGTGAGGTTACCAACCACTCTCCGATCACATC 1388
QY 341 LeuProGLInTYRLeuArgProValGLUASPValAlaThrSerGLInASPAspCYerTYR 360
DB 1389 CTTCCGACGACAAATCCGCGGCGCAGTGAAGATGTGCGCACGTCCCAAGACGACTGTTAC 1448
QY 361 LysPheAlaIleSerGLInserSerThrGLYThrValMetGLYAlaIleMetGLY 380
DB 1449 AAGTTGCACTCACTCAAGTCACTCAAGGACCTGTATGGAGCTGTATCATGAGAGGC 1508
QY 381 PheTYRValValPheASPArgAlaArgLYLeuArgIleGLYPhelAlaIleSerAlaCYeHIS 400
DB 1509 TTCTACGTTGTCTTGAATCGGGCCGAAAAGAAATGTGGCTTGTGTCAGAGCTTGCCAT 1568
QY 401 ValHISASPGLUPheArgThraAlaIleValAlaGLUGIYProPheValThrLeuASPMeGLU 420
DB 1569 GTGACACATGATGTTCAAGACCGGCGCGGTGAAGGCCCTTTTGTTCACCTTGACATGAA 1628
QY 421 ASPCYGLITYRAsnIleProGLInThraSPGLYSerThr 433
DB 1629 GACTGTGCTACAACTTCCACAGACAGATGATCAACC 1667
```

Search completed: August 22, 2005, 05:09:29  
Job time : 327 secs



Oy		181	VallVleGIlnrhHsValProamLeuPheSerLeuGlIneUCysGslYalaglYPhePro	200
Db		610	GTAAGACAGACCACGGTTCCTCCAACTCTCTCCCTGCAGCTGTGGTGCGCTTCCC	669
Oy		201	LeuaSngInserGIuValLeualaseValIGlYSerMetIleIleglYilLeap	220
Db		670	CTCAACACAGCTGAAGTGGCTGCTCtGTGCAGGGAGCATGATCATTTGGAGGTATCGAC	729
Oy		221	HieSerLeuYrThrGlsSerLeuTPyrTrhProIIeaRgsluTPyrTYrglu	240
Db		730	CACCTGGCTGACACAGCAGCTCTGTGTATCACCCATCCGGCGGAGTGTATTATGAG	789
Oy		241	VallIleIleValargValGIuIleaSnglYGIAspLeuLysMerApcCysLvsGIuTYr	260
Db		790	GTCATCATTTGTGGGGTGGAGATCAATGACAGAGACTGAAAANTGACTGCCAGAGTAC	849
Oy		261	ASrTYrAspLysSerIleValaspSerGIYrThrThraSnuArgLeubProIlySLysVal	280
Db		850	AACATAAGACAAGAAGCATTTGTGGCAGGTGGCACACCAACCATTCGTGGTCCCAAGAAAGTG	909
Oy		281	PhegluAlaalaValLysSerIleUysAlaIaseSerThGIuLysPheProaspGIY	300
Db		910	TTTGAAGCTGCAGTCAATTCATCAAGGACGCTCTCCACGAGAGTAGTCCCTGATGGT	969
Oy		301	PHeTPrLeuGIYGIuGlInleuValCYstrPGlnalagLYrThrProTrpaSniIephe	320
Db		970	TTCGTGCTGGAGAGAGAGCTGTGTGTCTGGCAAGACAGGACACACCCTTGGAACATTTTC	1022
Oy		321	ProValIlaSerLeuTYrleuMeTcGIYGIuValThraSngInserPheArgIleThrIle	340
Db		1030	CCAAGTATCTCACTCTACTTAATGGGTAGGTATACCAACAGTCTTCGCAATCACCATTC	1083
Oy		341	LeuPProGIInGIYrleuArgProValGIuaspValaIarhSerGIuaspAPCystrYr	360
Db		1090	CTTCCGACGACATATACCTGGCGCAATGGAGAAATGTGGCACGATCCCAAGACAGACTGTTC	1149
Oy		361	LysPheAlaIleSerGIInSerSerThrGIYrThrValMetGIYalavalIleMetGIuGLY	380
Db		1150	AA GTTGGCATCTCAAGTCAATCCACGGGCACTGTATGGAGAGCTGTATCATGAGGGGC	1209
Oy		381	PheTYrValValPheaspARGalaaRgLYsaRgIleGIYpheAlaValaserAlaCYshIs	400
Db		1210	TTCTACGTGTCTTTGATCGGGCCGGAABAGAAATGGCTTTGCTGTCAAGCGCTTGGCAT	1268
Oy		401	ValHisaspGIuPheArgThrAlaAlaValIGluyProPheValThrLeuaspMetGIu	420
Db		1270	GTGACAGATGAGTTCAGGACGGCAGCGGTGGAAAGCCCTTTTGTACCTTGACATGGAA	1329
Oy		421	ASPcYsgIYrAsnIleProGIInThrapGIusErThr	433
Db		1330	GACTGTGGCTAACATTCACACAGACAGTATAGTCAACC	1368
<b>RESULT 2</b>				
ID	AAA28278		standard; cDNA; 1503 BP.	
AC	AAA28278;			
XX				
DT	12-FEB-2001	(first entry)		
XX				
DE			Human cDNA encoding beta-secretase.	
KX			Beta-secretase; enzyme; amyloid plaque; Alzheimer's disease; human;	
KW			Down's syndrome; amyloid angiopathy; gene therapy; neuroprotective; ss.	
XX				
OS			Homo sapiens.	
FH	Key		Location/Qualifiers	
FT	CD5		I..1503	
FT			/*tag= a	
FT			/product= "Beta secretase"	
FT			/note= "No from codon given"	



Db 1324 GACTGTGGCTTCAACATTCCACAGACAGATGATCAACC 1362

RESULT 3  
ID AAA59550  
AAA59550 standard; DNA; 1503 BP.  
XX  
AC AAA59550;  
XX  
DT 14-NOV-2000 (first entry)  
XX  
DE DNA encoding a human beta-secretase enzyme.  
XX  
KM Beta-secretase; beta-amyloid precursor protein; beta-amyloid peptide;  
KM amyloid plaque component; Alzheimer's disease; amyloidogenic disease;  
KM inhibitor; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..1503  
FT /tag=a  
FT /product="beta-secretase"  
FT /note="no termination codon given"  
XX  
PN MO200047618-A2.  
XX  
PD 17-AUG-2000.  
XX  
PF 10-FEB-2000; 2000MO-US003819.  
XX  
PR 10-FEB-1999; 99US-0119571P.  
PR 15-JUN-1999; 99US-0139172P.  
XX  
PA (ELAN-) ELAN PHARM INC.  
XX  
PI Anderson JP, Basi G, Doane MT, Frigon N, John V, Power M,  
PI Sinha S, Tateuno G, Tung J, Wang S, Mcconlogue L,  
XX  
DR WPI; 2000-533011/48.  
XX  
P-PSDB; AAB07896.  
XX  
PT Purified beta-secretase protein used in assays to discover inhibitors  
PT which can be used for the treatment of amyloidogenic diseases e.g.  
PT Alzheimer's disease.  
XX  
PS Disclosure; Fig 1A; 121pp; English.  
XX  
CC The specification describes a beta-secretase enzyme. The enzyme cleaves  
CC beta-amyloid precursor protein to produce beta-amyloid peptide. This  
CC enzyme is therefore implicated in the production of amyloid plaque  
CC components which accumulate in the brains of individuals afflicted with  
CC Alzheimer's disease. Inhibitors of beta-secretase are administered to a  
CC mammalian subject e.g. with Alzheimer's disease or Alzheimer's disease-  
CC like pathology to test if they maintain or improve cognitive ability or  
CC reduce the plaque burden. The compounds are used for the treatment of  
CC amyloidogenic diseases e.g. Alzheimer's disease. The present sequence  
CC encodes a human beta-secretase enzyme  
XX  
SQ Sequence 1503 BP; 305 A; 448 C; 431 G; 319 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 2,046-227 Length: 1503  
Score: 2267.00 Matches: 429  
Percent Similarity: 99.31% Conservative: 1  
Best Local Similarity: 99.08% Mismatches: 3  
Query Match: 99.08% Indels: 0  
Gaps: 0

US-10-726-967a-84 (1-433) x AAA59550 (1-1503)

QY 1 ThrGlnHisGlyIleArgLeuProLeuArgSerGlyLeuGlyValAlaProLeuGlyLeu 20  
|||||

Db 64 ACCGAGCAGCGGACCTCGGCTGCCAGCGGCTGGGGGGGCCCCCTGGGGCTG 123

QY 21 GluIleAsnLeuGlyThrAspGluGluProGluGluProGlyIleArgAspGlySerPheVal 40  
:::

Db 124 CGGCTGCCCGGGAGACCGAGCAGAGAGCCCGAGAGCCCGGAGGGGCGAGCTTTGTG 183

QY 41 GluMetValAspAsnLeuArgGlyIleSerGlyGlnGlyIleValGluMetThrVal 60

Db 184 GAGATGGTGGACCAACTGAGGGGCAAGTCGGGGGCGAGGCTACTACGTGAGATGACCGTGG 243

QY 61 GlySerProProGlnThrLeuAsnIleLeuValAspThrGlySerSerAsnPheAlaVal 80

Db 244 GCGAGCCCCCGCAGACCTCAACATCTCGTGGATACAGGACAGCACTTTGCGAGTG 303

QY 81 GlyAlaAlaProHisProPheLeuHisArgIleValGlnArgGlnLeuSerSerThrTyr 100

Db 304 GGTGTGCTGCCCCCAACCCCTTCTCGATGCGTCACTACAGAGGAGCGTGTCCAGACATAC 363

QY 101 ArgAspLeuArgGlyGlyValIleValProTyrThrGlnGlyIleValGluLeu 120

Db 364 CGGAGCCTCCGGAGGGGTGTGTATGTGCCCTTACACCCAGGGGCAAGTGGAGGGAGCTGG 423

QY 121 GlyThrAspLeuValSerIleProHisGlyProAsnValThrValArgAlaAsnIleAla 140

Db 424 GGCACCGACCTGTGACGATCCCGCATGGCCCAACGTCACTGTGCGTCCCAACATTTGCT 483

QY 141 AlaIleThrGluSerAspIlePhePheIleAsnGlySerAsnTrpGluGlyIleLeuGly 160

Db 484 GCCATCACTGAAATCAGACAAAGTTCTTCATCAACGGGCTCAACATGGAGGCACTCTGGG 543

QY 161 LeuAlaTyrAlaGluIleAlaArgProAspAspSerLeuGluProPheAspSerLeu 180

Db 544 CTGGCTATGTGAGATTCAGGCTCAGGCTCAGACCTCCCTGAGGCTTTTGTGACTCTCTG 603

QY 181 ValIleGlnThrHisValProAsnLeuPheSerLeuGlnLeuGlyAlaGlyPhePro 200

Db 604 GTAAAGCAGACCCACGTTCCCAACCTCTCTCCCTGAGCTTTGTGTGCTGCTTCCCC 663

QY 201 LeuAsnGlnSerGluValIleAsnIleSerValGlyGlySerMetIleIleGlyIleAsp 220

Db 664 CTCAACGAGCTGAAGTCTGGGCTCTGTGCGAGGAGCATGATCATTTGAGGATTCGAC 723

QY 221 HisSerLeuTyrThrGlySerLeuTyrTyrThrProIleArgArgGluTyrTyrArgIu 240

Db 724 CACTCGCTGACAGGACGATCTCTGATACCCATCCGCGGAGGTGATTTATGAG 783

QY 241 ValIleIleValAlaArgValGluIleAsnGlyGlnAspLeuIleMetAspCysIleGlyTyr 260

Db 784 GTGATCATTTGTGGGGTGGAGATCAATGACAGATCTGMAAATGATCGCAAGAGATTC 843

QY 261 AsnTyrAspIleSerIleValAspSerGlyThrThrAsnLeuArgLeuProIleValVal 280

Db 844 AACCTAGACAAAGCATTTGGACAGTGGACACCAACCTTCTGTTGGCCAAAGAAAGTG 903

QY 281 PheGluAlaAlaValIleSerIleIleValAlaSerSerThrGluIlePheProAspGly 300

Db 904 TTTGAGGCTGCAATCAATCCATCAAGGACGCTCTCCACGAGAAAGTTCCCTGATGAT 963

QY 301 PheTrpLeuGlyGluGlnLeuValCysTrpGlnAlaGlyIleThrThrProTrpAsnIlePhe 320

Db 964 TTCTGCTAGAGAGACGCTGTGTGCTGGCAAGCAGCACCCTTGGAACATTTTTC 1023

QY 321 ProValIleSerLeuTyrLeuMetGlyValIleThrAsnGlnSerPheArgIleThrIle 340

Db 1024 CCGATATCTCACTTCACTTAATGGGTGAGTTTACCAACAACTCTTCCGCAATCCATC 1083

QY 341 LeuProGlnGlnTyrLeuArgProValGluAspValAlaThrSerGlnAspAspCysTyr 360

Db 1084 CTTCCGACCAATATACCTGGCGGCAGTGGAAGATGTGGCAGCTCCCAAGACACATCTTAC 1143

QY 361 IysPheAlaIleSerGlnSerSerThrGlyThrValMetGlyAlaValIleMetGluGly 380

Db 1144 AAGTTTGCATCTCAACATCATCCAGGGCAGCTTTATGGAGCTGTATCATGTGAGGGGC 1203

Oy		381	pheTYrValIvalPheaSprxglaAArglySAryllEglYPheaIalValSerilaCySHis	400
Db		1204	TTTACGGTGTGCTTGATCGGCGCCGAANAAGATTGGCTTCCTGTCAAGCCTTGCCAT	126
Oy		401	ValHisAPrglUPheaYrGThraIalaValaIgUglYProPheValThrLeuAPMetGlu	420
Db		1264	GTCGACGATGACTTCACGACGGGAGCGGTGGAAAGGCCCTTTGTGCACCTTGACATGGAA	1322
Oy		421	AapCYsgLYTYrAsnIlleProGlnThrAapGIuSerThr	433
Db		1324	GACTGTGGCTTACACAATTCACACAGACATGATGATAACC	1362
RESULT 4				
ACC84849			standard; DNA; 1506 BP.	
XX	AC	ACC84849;		
XX	DT	12-SEP-2003	(first entry)	
XX			Human memapsin 2 protein encoding DNA.	
XX	DE			
XX	KW		Memapsin 1; neurotrophic; neuroprotective; memapsin 2; beta secretase;	
KW			beta-amyloid protein; Alzheimer's disease; human; gene; ds.	
OS			Homo sapiens.	
XX	FH	Key	Location/Qualifiers	
FH	CDS		1..1506	
FT			/tag= a	
FT			/product= "memapsin 2"	
FT	eig_peptide		1..63	
FT			/tag= b	
FT	mat_peptide		64..1503	
FT			/tag= c	
XX				
PN		WO2003039454-A2.		
PD		15-MAY-2003.		
PR		23-OCT-2002; 2002MO-US034324.		
XX				
PR		23-OCT-2001; 2001US-0335952P.		
PR		27-NOV-2001; 2001US-0333545P.		
PR		14-JAN-2002; 2002US-0348464P.		
PR		14-JAN-2002; 2002US-0348615P.		
PR		20-JUN-2002; 2002US-0390804P.		
PR		19-JUL-2002; 2002US-0397557P.		
PR		19-JUL-2002; 2002US-0397619P.		
PA		(OKLA-) OKLAHOMA MEDICAL RES FOUND.		
PA		(UNIT ) UNIT ILLINOIS FOUND.		
Pi		Shosh AK, Tang J, Bilcer G, Chang W, Hong L, Koelsch G, Loy J;		
Pi		Turner RT;		
DR		WPI: 2003-541410/51.		
XX		P-PADB; ABR61928.		
PT		New peptide compounds are memapsin beta secretase inhibitors used for		
PT		treating Alzheimer's disease.		
PS		Claim 101; Fig 8; 407pp; English.		
XX		The invention relates to peptid compounds of specified formula. The		
CC		compounds exhibit memapsin 2-beta secretase inhibitory activity relative		
CC		to memapsin 1-beta secretase and reduce the accumulation of beta-amyloid		
CC		protein. The compounds can be used for treating Alzheimer's disease. The		
CC		present sequence represents a human memapsin 2 protein encoding DNA		
CC		(Genbank Index (GI :21040369))		
SQ		-Sequence 1506 BP; 306 A; 449 C; 431 G; 320 T; 0 U; 0 Other;		

Alignment Scores:	2.05e-227	Length:	1506
Pred. No.:	2267.00	Matches:	429
Score:	99.31%	Conservative:	1
Percent Similarity:	99.08%	Mismatches:	3
Best Local Similarity:	99.08%	Indels:	0
Query Match:	9	Gaps:	0
US-10-726-967A-84 (1-433) X ACC84849 (1-1506)			
QY	1 ThrGlnHisGlyIleArgLeuProLeuArgSerGlyLeuGlyValAlaProLeuGlyLeu	20	
DB	64 ACCGACGACGGCAGTCCGGCTGCGCAGCGGCGCGGGGGGGGGCCCCCTGGGGCTG	123	
QY	21 GluIleAsnLeuGluThrAspGluGluProGluGluProGlyArgArgGlySerPheVal	40	
DB	124 CGCGTGGCCCCGGGAGACCGACGAGAGGCCCGAGAGGCCCGCGGAGGGGCGAGCTTGTG	183	
QY	41 GluMetValAspAsnLeuArgGlyLysSerGlyGlyGlyIleValAlaLeuMetThrVal	60	
DB	184 GAGATGGTGGACCACTGAGGGGCGAGTGGGGCGAGGCGCTTACGTGAGATGACCGTG	243	
QY	61 GlySerProProGlnThrLeuAsnIleLeuValAspThrGlySerSerAsnIleVal	80	
DB	244 GGCAGCCCCCGGACAGCCTCAACATCTGGTGGATACGAGGACGAGCATCTTGGAGTG	303	
QY	81 GlyAlaAlaProHisProPheLeuHisArgTyrTyrGlnArgGluLeuSerSerThrTyr	100	
DB	304 GGTGTGCCCCCGACCCCTTCTGCAATGCTACTACAGAGGACAGTGTCCAGACATAC	363	
QY	101 ArgAspLeuArgLysGlyValTyrValProTyrThrGlyLysTyrPgluGlyLeu	120	
DB	364 CGGACCTCCGGAGAGGTGTATGTGCCCTACACCGAGGGCGAGTGGAGGGGAGCTG	423	
QY	121 GlyThrAspLeuValSerIleProHisGlyProAsnValThrValArgAlaAsnIleAla	140	
DB	424 GGCACCGACCTGGTAGAGATCCCGCAAGGCCCGACAGTCACTGTGTGTCCACATTTGCT	483	
QY	141 AlaIleThrGluSerAspLysPhePheIleAsnGlySerAsnTrpGluGlyIleLeuGly	160	
DB	484 GCCATCATCTGANTCAGACAGATTTCTTCAACGGCTCCAACTGGGAGAGGCAATCCGGGG	543	
QY	161 LeuAlaTyrAlaGluIleAlaArgProAspAspSerLeuGluProPhePheAspSerLeu	180	
DB	544 CTGGCTATGTGAGATGTCAGAGCTCCAGGCGTCACTCCCTGAGAGCTTTCTTGTGACTCTG	603	
QY	181 ValLysGlnThrHisValProAsnLeuPheSerLeuGluLeuCyseGlyValGlyPhePro	200	
DB	604 GTAAAGCGAGACCCAGCTTCCAACTCTTCCCTGCGACCTTGTGTGTGTGCTTCC	663	
QY	201 LeuAsnGlnSerGlyValLeuAlaSerValGlyGlySerMetIleIleGlyGlyIleAsp	220	
DB	664 CTCAACACAGCTGAGAGTGTGGCTCTGTGTGCGAGGAGCATGATCATTTGAGGTATGAG	723	
QY	221 HisSerLeuTyrThrGlySerLeuTyrTyrThrProIleArgArgGluTyrTyrGlu	240	
DB	724 CACTGGCTGTACACAGGCGAGTCTCTGTGTATACACCATCCGGGCGGAGTGTATTATGAG	783	
QY	241 ValIleIleValArgValGluIleAsnGlyGlnAspLeuLysPheSerCyseLysGlyTyr	260	
DB	784 GTCATCATTTGTGGGTGGAGATTCATAGGACAGGATCTAAATAATGAGCTGCAAGAGTAC	843	
QY	261 AsnTyrAspLysSerIleValAspSerGlyTyrThrAsnLeuArgLeuProLysVal	280	
DB	844 AACTATGACAGAGCATTTGTGACAGTGGCACCAACCACTTCGTTGGCCAGAAAGTG	903	
QY	281 PheGluAlaAlaValLysSerIleLysAlaIleSerSerThrGluLysPheProAspGly	300	
DB	904 TTGGAAGCTGAGTCAATTCATCAAGGCGAGCTCTCCACGAGGAAGTTCTCTGTATGCT	963	
QY	301 PheTyrPheGlyGluGluMetValCyseTyrGlnAlaGlyThrThrProTyrAsnIlePhe	320	

Db	964	TTCTGGCTAGAGAGACAGCTGGTGTCTGCGCAAGCAGGACCAACCCCTTGGAACATTTTC	1023
Qy	321	ProValIILSerLeuTyrLeuMetGlyIValIThrAsnGlnSerPheArgIleThrIle	340
Db	1024	CCAGTCATCTCACTTACCTTAATGGGTGAGGTATACCAACGATCCTTCGCGATCACATTC	1083
Qy	341	LeuProGlnGlyTyrLeuAspProValIGluAspValAlaThrSerGlnAspAspCysTyr	360
Db	1084	CTTCCGACGACATACCTCGGCGCAGGGAAGATGGGCCACATCCCAAGACGACCTGTAC	1143
Qy	361	LysPheAlaIILSerGlnSerSerThrGlyThrValMetGlyAlaValIILewetGluGly	380
Db	1144	AAGTTGGCATCTCAACAGTCATCCACGGGCGGACCTGTTATGGAGCGTGTATCATGAGAGGC	1203
Qy	381	PheTyrValValPheAspArgAlaArgIyValArgIleGlyPheAlaValISerAlaCysHis	400
Db	1204	TTCTACGTTGTCTTGATGGGCGCCGAAACCAATTTGGCTTGTCTGACGCGCTTGGCAT	1263
Qy	401	ValIILAspGluPheArgThrAlaAlaValIGluGlyProPheValThrLeuAspMetGlu	420
Db	1264	GGCGACGATGAATTCAGACGCGCAGCGGTGGAAAGCCCTTTTGTCACTTGACATGGAA	1323
Qy	421	AspCyGleGlyTyrAsnIILProGlnThrAspGluSerThr	433
Db	1324	GACTGTGGCTACACATTCACAGACAGATGATGCACCC	1362
RESULT 5			
ADL18183			
ID	ADL18183	standard; cDNA; 1506 BP.	
XX	ADL18183;		
XX	06-MAY-2004 (first entry)		
XX	Human APP beta-secretase encoding cDNA SEQ ID NO:103.		
DE	chimeric protein; signal protein; trafficking signal targeting;		
KW	proteolytic cleavage site; protease; protease inhibitor; enzyme; human;		
KM	APP beta-secretase; gene; ss.		
XX			
OS	Homo sapiens.		
XX			
PN	MO2003014381-AL.		
PD	20-FEB-2003.		
XX			
PF	08-AUG-2002; 2002WO-KR001515.		
XX			
PR	10-AUG-2001; 2001KR-00048123.		
XX			
PA	(AHRA-) AHRAM BIOSYSTEMS INC.		
XX			
PI	Hwang I, Kim DH, Lee YJ;		
XX	WPI; 2003-256596/25.		
DR	P-PSDB; ADL18184.		
PT	New chimeric protein, useful for detecting protease inhibitors inside the		
XX	cell or tissue.		
XX			
XX	Disclosure; SEQ ID NO 103; 214pp; English.		
XX			
XX	The present invention describes a chimeric protein comprising at least		
CC	one signal protein that has a trafficking signal targeting to a		
CC	subcellular organelle and at least one proteolytic cleavage site for a		
CC	protease. The chimeric protein is constructed, so that: (a) the		
CC	trafficking signals of all the signal proteins are inactivated by linking		
CC	the proteolytic site or a signal masking protein through the proteolytic		
CC	site to the N- or C- terminus of the signal proteins, and so the chimeric		
CC	protein is present in cytosol; (b) the trafficking signal of at least one		
CC	signal protein is activated when the proteolytic cleavage site is cleaved		
CC	by the protease, and as a result at least one fragment protein that		
CC	includes the activated signal protein is a transported to a subcellular		

organelle; and (c) the chimeric protein is labelled with at least one fluorescent protein and the position and intensity distribution of the fluorescent label signal in the cell is altered depending on the cleavage by the protease. Also described: (1) a recombinant gene comprising a nucleic acid sequence encoding the chimeric protein which is constructed to express the chimeric protein in a cell; (2) a cell transformed with the recombinant gene or vector; (3) analysing the activity of a protease *in vivo*; (4) screening protease inhibitors *in vivo*; (5) a system for detecting a protease inside a cell; (6) a nucleic acid comprising the sequence encoding the chimeric protein for detecting protease activity in a cell; (7) a vector comprising the nucleic acid; (8) a kit for detecting a protease inside a cell comprising the chimeric protein or the vector; (9) detecting a protease inside a cell or tissue; and (10) detecting a protease inhibitor *in vivo*. The chimeric protein is useful for detecting protease inhibitors inside the cell or tissue. The present sequence encodes a human APP beta-secretase, which is used in the exemplification of the present invention.

SQ Sequence 1506 BP; 306 A; 449 C; 431 G; 320 T; 0 U; 0 Other;

**Alignment Scores:**

Fried. NO.:	2.035e-22/	Length:	1506
Score:	2267.00	Matches:	429

Percent Similarity:	99.31%	Conservative:	1
Best Local Similarity:	99.08%	Mismatches:	3

Query Match:	99.08%	Indels:	0
DB:	10	Cons:	0

TFC-10-536 OCT22 04 (1 433) = 50710103 (1 166)

[illegible][illegible]

0? ALLCAGACGCGAICGGCCIGCGCAAGCGCLGGGGG

zy z1 gni1eashLeugiuinraps6ugluProglugluProglYAT  
:::

Db 124 CGCTGCCCCGGAGACCGACGAAGAGCCCCGAGAGCCCCGGCCG

QY 41 GluMetValAspAsnLeuArgGlyLysSerGlyGlnGlyTyrTyr

Db 184 GAGATGTGACCACTGAGGGCAAGTCGGGCGAGGCTACTA

61 GlySerProProGlnThrLeuAsnIleLeuValAspThrGlySer

Db 244 GGCAGCCCCCGCAGACGCTCAACATCCTGTGTGATACAGGCAG

QY 81 GlyAlaAlaProHisProPheLeuHisArgTyrTrpGlnArgGln

Db 304 GATGCTGCCCCCAACCCCTTACATGCACTTACAGAGCA

101 ArcAsnIeuAraIwG] vVa] TuvvA] proTuvvthwG] nG] vi v

364

22

[illegible]

727 00CACCAGACCGTATGACAACCCCAAGCCCAACGCACACGG

141 AlAlenHrGruSerAplYsneHneIeaShGlySerAanTr

DB 484 GCCATTCACIGATCAGACCAAGTTCCTTCATCAACCGCTCCAACTG

161 LeuAlaTyrAlaGluIleAlaArgProAspPaspSerLeuGluPrr

Db 544 CTGGCCATGCTGAGATTGCCAGGCGCTGACGACTCCCTGGAGCC

181 VallysglNThrhlsValProAsnLeuPheSerLeuGlnLeuCy

Db 604 GTAAAGCAGACCCACGTTCCCAACTCTTCTCTCCCTGCAGCTTTC

201 LeuAsnGlnSerGluValLeuAlaSerValGlyGlySerMetIle

664 CTCAACCACTCTGAAGTGCCTGCCCCTCTGTCCGAGGGAGCATGAT



QY 221 HisSerLeuYrThGlySerLeuTrpYrThrProIleArgGluTrpYrTrGlu 240  
DB CACTCGCTGTACAGCAGGAGTCTGTGTATACACCCATCCGGCGGAGTGTATTATGAG 783  
QY 241 ValIleIleValArgValGluIleAsnGlyGlnAspLeuValMetAspCysLeuGlyTr 260  
DB GTCAATCATTTGCGGGTGAAGATCAATGACAGATCGAATGAAATGACATGCAAGAGTAC 843  
QY 261 AsnYrAspLysSerIleValAspSerGlyThrThrAsnLeuArgLeuProLysVal 280  
DB AACTATGACAGAGCATTTGTGACAGTGGCACCCACCACTTCCTTTCGCCAAGAAAGTG 903  
QY 281 PheGluValAlaValIleSerIleLysValAlaSerSerThrGluLysPheProAspGly 300  
DB TTTGAAAGCTGACGATCAAACTTCATCAAGGACGCTCTCCACGAGAAAGTTCCCTCATG 963  
QY 301 PheTrpLeuGlyGluGlnLeuValCysTrpGlnAlaGlyThrThrProTrpAsnIlePhe 320  
DB TTCTGGCTAGAGACAGCAGCTGGTGTGCTGGCAAGGACCACTTCCTGGAAACATTTTC 1023  
QY 321 ProValIleSerLeuYrLeuMetGlyGluValThrAsnGlnSerPheArgIleThrIle 340  
DB CCAATCATCTCACTTCACTTAATGGGTGAGGTTACCAACCACTCTTCGCCATCACATC 1083  
QY 341 LeuProGlnGlnYrLeuArgProValGluAspValAlaThrSerGlnAspAspCysTr 360  
DB CTTCGCGACGACATACCTGGCGCAGTGAAGATGTGGCCACGTCACAGACATCTGTAC 1143  
QY 361 LysPheAlaIleSerGlnSerSerThrGlyThrValMetGlyValIleMetGluGly 380  
DB AAGTTTGCATCTCAACATTCACATTCACGCGGCTGTATGAGGAGCTGTATATCAATGA 1203  
QY 381 PheYrValValPheAspArgAlaArgLysArgIleGlyPheAlaValSerAlaCysHis 400  
DB TTCTAGCTTGTCTTGTATCGGCGCCGAAACCAATGGCTTGTCTGCTGACGCTTGGCAT 1263  
QY 401 ValHisAspGluPheArgThrAlaAlaValGluGlyProPheValThrLeuAspMetGlu 420  
DB GTGCAAGATGATTCAGACGCGCAGCGGTGGAAGGCCCTTTGTCACTTGGACATGGA 1323  
QY 421 AspCysGlyYrAsnIleProGlnThrAspGluSerThr 433  
DB 1324 GACTGTGGCTACAACTTCACAGACAGATGATCAACC 1362

RESULT 6  
AD082259  
ID AD082259 standard; cDNA; 1524 BP.  
XX  
XX AD082259;  
XX  
XX 21-OCT-2004 (first entry)  
XX  
XX Human BACE1 coding sequence.  
XX  
XX transgenic animal; beta-site amyloid precursor protein cleaving enzyme;  
XX BACE1; Hemostatic; Neuroprotective; Neotropic; Beta-secretase inhibitor;  
XX amyloid beta; neurodegenerative disease; Alzheimer's disease;  
XX cerebral amyloid angiopathy; Lewy body dementia; Down's syndrome;  
XX hereditary cerebral hemorrhage; amyloidosis; Guam Parkinson-Dementia; SS.  
XX  
XX Homo sapiens.  
XX  
XX Synthetic.  
XX  
XX MO2004062627-A2.  
XX  
XX 29-JUL-2004.  
XX  
XX 13-JAN-2004; 2004MO-US000883.  
XX  
XX 13-JAN-2003; 2003US-0439633P.  
XX  
XX (REGC ) UNIV CALIFORNIA.  
PA

XX  
PI Maslian E, Rockenstein E;  
XX  
XX WPI; 2004-544036/52.  
DR  
XX  
XX New non-human transgenic animal overexpressing the (human) beta-site  
PT amyloid precursor protein cleaving enzyme (BACE1), useful as a model for  
PT conditions such as Alzheimer's disease, and in screening for therapeutic  
PT agents.  
XX  
XX Example 1; SEQ ID NO 1; 54pp; English.  
XX  
XX The present invention relates to a transgenic non-human animal for  
CC overexpressing beta-site amyloid precursor protein cleaving enzyme  
CC (BACE1) comprising cells containing a DNA transgene encoding for BACE1.  
CC The transgenic non-human animal is useful as a disease model, in studying  
CC the in vivo and in vitro regulation and effects of BACE1 in specific  
CC tissue types, in examining the role of BACE1 proteins in the accumulation  
CC of amyloid beta, and for developing therapies for amyloid beta-related  
CC conditions. The method, agents or compositions are useful for treating  
CC neurodegenerative disease, e.g. Alzheimer's disease, cerebral amyloid  
CC angiopathy, Lewy body dementia, Down's syndrome, hereditary cerebral  
CC hemorrhage with amyloidosis (Dutch type), or Guam Parkinson-Dementia  
CC complex. The present sequence represents human BACE1 coding sequence.  
XX  
SQ Sequence 1524 BP; 309 A; 456 C; 437 G; 322 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 2,086-227 Length: 1524  
Score: 2267.00 Matches: 429  
Percent Similarity: 99.31% Conservative: 1  
Best Local Similarity: 99.08% Mismatches: 3  
Query Match: 99.08% Indels: 0  
DB: 13 Gaps: 0  
US-10-726-967A-84 (1-433) x AD082259 (1-1524)  
QY 1 ThrGlnHisGlyTLeuArgLeuProLeuArgSerGlyLeuGlyValAlaProLeuGlyLeu 20  
DB 76 ACCCGACAGCGGATCGGCTGCTGCGCAGCGGCTGGGGGCGCCCTCGGGGCTG 135  
QY 21 GluIleAsnLeuGluThrAspGluGluProGluGluProGlyArgArgGlySerPheVal 40  
DB CGGCTGCCCCCGGAGAACCGACGAAAGGCCGAGAGCCCGCGGAGGGGCGCTTTGTG 195  
QY 41 GluMetValAspAsnLeuArgGlyLysSerGlyGlnGlyYrYrValGluMetThrVal 60  
DB GAGATGGTGGACAACTGAGGGGCAAGTCGGGGCAGGGCTACTACGTGAGATGACCGTG 255  
QY 61 GlySerProProGlnThrLeuAsnIleLeuValAspThrGlySerSerAsnPheAlaVal 80  
DB GGCAGCCCCCGCCACACGCTCAACATCTGTGTGATGACCCAGGCGCAAGTGGGAAAGGGAGCTG 315  
QY 81 GlyAlaAlaProHisProPheLeuHisArgYrYrGlnArgGlnLeuSerSerThrYr 100  
DB GGTGCTGCCCCCACCACCTTCTGTGATCGTACACAGAGCACTGTCCAGCACATTC 375  
QY 101 ArgAspLeuArgLysGlyValYrValProYrThrGlnGlyLysTrpGluGluLeu 120  
DB CGGGAACCTCGGAAGGGTGTATGTGCTTACACCCAGGCGCAAGTGGGAAAGGGAGCTG 435  
QY 121 GlyThrAspLeuValSerIleProHisGlyProAsnValThrValArgAlaAsnIleAla 140  
DB GGCACCGACCTGTGTAGACATCCCAATGCCCCCAAGTCACTGTGGTGCCAACTGTCT 495  
QY 141 AlaIleThrGluSerAspLysPhePheIleAsnGlySerAsnTrpGluGlyIleLeuGly 160  
DB GCATCATCTGAATTCAGACAAAGTTCTTCATCAACGGCTCCAATGGGAAGCATCTCGGG 555  
QY 161 LeuAlaYrAlaGluIleAlaArgProAspAspSerLeuGluProPheAspSerLeu 180  
DB CTGGCCTATGCTGAGATTGCGAGGCTGACGACTCCCTGAGACCTTTCTTACTCTCTG 615

	Accession	Gene	Protein	Location/Qualifiers
QY	181	VallyegIntHrHleVal	ProAbnLeuPheSerLeuGlnLeuCySGlyAlaGlyPhePro	200
Db	616	GTAAGCGAGACCACGTTCCCACTCTTCTCCCTGCGAGCTTGTGGGCTTCC		675
QY	201	LeuengInserGluValLeuAlaSerValGlyGlySerMetIleIleGlyValLeaP		220
Db	676	CTCAACCAAGTCTGAAGAGCTGGCTCTGTCCGAGAGGAGCAATGATGAGGATTCGAC		735
QY	221	HisSerLeuTYrThrGlySerLeuTrpTYrThrProIleArgArgIuTPYrTYrGlu		240
Db	736	CACTCGCTGATCACAGGACAGCTCTGGATATACCACTCCGGCGAGTGATATATAG		795
QY	241	ValIleIleValaGValGluIleAsnGlyGlnAspLeuYsMetApyCylySGlyTYr		260
Db	796	GTGATCATGTGCGGGTGGAGATCAAGAGACAGATCTGAAATATGACTCAGAGAGTAC		855
QY	261	AsnTYrAspLYsSerIleValaAspSerGlyTYrThrAsnLeuArgLeuPolslyVal		280
Db	856	AACATATACAAAGACATTGTGACAGTGGCACCAACCACTTGTTGGCCAAAGAAAG		915
QY	281	PheGluAlaAlaValLYsSerIleLYsAlaIleSerSerThrGluYsPheProAspGly		300
Db	916	TTTGAACTCAGTCGCAAAATCATCAAGGACGCTCTCCACGGAAGAAAGTTCCCTAGT		975
QY	301	PheTrpLeuGlyGluGlnLeuValCysTrpGlnAlaGlyTYrThrProTrpAsnIlePhe		320
Db	976	TTCCTGCTGAGAGACAGCTGAGTGTGTGTGGCAAGCAGCACCACTTGGAAACATTTTC		1031
QY	321	ProValIleSerLeuTYrLeuMetGlyGluValTYrAsnGlnSerPheArgIleThrIle		340
Db	1036	CCAGTCATCTCACTTAACCTTAATGGGTGAGGTTACCAACAGTCTTCCGATCACATTC		1095
QY	341	LeuProGlnIleTYrIleuArgProValGluAspValAlaThrSerGlnAspApyCysTYr		360
Db	1096	CTTCCGAGCAAAATCTGGCGGCCAGTGGAAAGATGGGCCACGCTCCCAAGACACATGTTAC		1155
QY	361	LYsPheAlaIleSerGlnSerSerThrGlyTYrValMetGlyAlaValIleMetGluGly		380
Db	1156	AAGTTGGCCATCTCAAGTCATCCACGGGCACTGTTATGGAGAGCTTTATCATGAGGGC		1211
QY	381	PheTYrValValPheAspArgAlaArgLYsValGlyIleGlyPheAlaValSerAlaCysHis		400
Db	1216	TTCCTAGTGTCTTGTGATGGGGCCGAAAAACAATTTGGCTTGTGTGACGCTTGCCAT		1275
QY	401	ValHisAspGluPheArgThrAlaValaValGluGlyProPheValThrLeuAspMetGlu		420
Db	1276	GTGACAGAGATTCACAGACGCGAGGGTGAAGGCCCTTTTGTCACTTGACATGGAA		1335
QY	421	AspCYsGlyTYrAsnIleProGlnThrAspGlnSerThr		433
Db	1336	GACTGTGGCTACACATTCACACAGACAGATGAGTCAACC		1374
RESULT 7				
ID	ABA02406	standard, cDNA, 1527 BP.		
XX	ABA02406;			
AC	26-FEB-2002	(first entry)		
DT	XX			
XX	XX			
DE	XX	FLAG-tagged human beta-secretase encoding cDNA.		
XX	XX			
KM	XX	Human, beta-secretase; FLAG tag; inhibitor; amine compound;		
KM	XX	beta amyloid protein production; head injury; spinal injury;		
KM	XX	amyloid precursor protein alpha secretion; nerve damage;		
KM	XX	meningitis sequelae; cerebral paralysis; memory disorder; mental disease;		
KM	XX	nootropic; neuroprotective; cerebroprotective; ss.		
OS	XX	Homo sapiens.		
OS	XX	Synthetic.		
Key	XX	Location/Qualifiers		
FT	XX	1..1527		
FT	XX	CDS		

FT	/tag= a
FT	/partial
FT	/product= "FLAG-tagged human beta-secretase"
FT	/note= "No stop codon given in the specification"
XX	
PN	W0200187293-A1.
PD	22-NOV-2001.
XX	
PF	18-MAY-2001, 2001WO-JP004144.
XX	
PR	19-MAY-2000, 2000JP-00152758.
XX	
PA	(TAKE ) TAKEDA CHEM IND LTD.
XX	
P1	Miyamoto M, Matsui J, Fukumoto H, Tani N;
XX	
DR	WPI; 2002-055640/07.
XX	P-PDB; AAMS2697.
PT	Beta-secretase inhibitor used for treating e.g. Alzheimer's disease and
XX	injury to brain or spine, and neurodegeneration, comprising amine
PT	compound.
XX	
PS	Example; Page 78-79; 86pp; Japanese.
XX	
CC	The invention relates to novel amine compounds which are beta-secretase
CC	inhibitors. The beta-secretase compounds also have the ability to promote
CC	amyloid precursor protein alpha secretion and to inhibit beta amyloid
CC	protein production. The beta-secretase inhibitors of the invention can b
CC	used for treating head or spinal injuries, nerve damage, sequelae of
CC	meningitis, cerebral paralysis, memory disorders and mental diseases. Th
CC	present sequence represents cDNA encoding a FLAG-tagged human beta-
XX	secretase used in the exemplifications of the invention
XX	
SQ	Sequence 1527 BP; 315 A; 451 C; 438 G; 323 T; 0 U; 0 Other;
Alignment Scores:	
Pred. No.:	2,09e-227
Score:	2267.00
Best Local Similarity:	99.31%
Percent Similarity:	99.08%
Query Match:	99.08%
DB:	6
	Gaps: 0
US-10-726-967A-84 (1-433) x ABA02406 (1-1527)	
QY	1 Thrglnhngllyleahngleuproleuargseglyleglvalproleuglyleu 20
Db	64 ACCGAGCAGCGACATCGGCTGCCCTGCCAGCGGCTGGGGGGGCCCTGGGGCTG 12
QY	21 Gluileanleugluthraapglugluproglugluproglylvararglyserpheval 40
Db	124 CGGCTGCCCGGAGACCGACGAGAGCCCGAGAGCCCGCCGAGGGGACGCTTTGTG 18
QY	41 Glumetvalaspsanleuarglylyserglynglyltyrtyvalglumetthrval 60
Db	184 GAGATGGTGCACCACTGAGGGGCAAGTCGGGGGAGGCTCTACTGAGATGACCGTG 24
QY	61 Glyserproproglinhrleuanleuvalapbrrnglyseraserasphalaval 80
Db	244 GCGAGCCCCCGAGACGCTCAACATCTGGTGGATACGGGAGCATTAACCTTGCAGTG 30
QY	81 Glylvalalaprohleprouhleuhlbarglytyrtyrjnatrglnleusersethrtyr 100
Db	304 GGTGTGCCCCCACCCTTCTTGATCGTACTACAGAGGACGCTGTCCAGCACATAC 36
QY	101 ArgAspLeuArgLySglValTyrtalprotyrthrnglnglylytprglnglygluleu 120
Db	364 CGGACCTCCGAGAGGATGTGTATGTGCTTACCCAGGGGCAAGTGGAGGGGACCTG 42
QY	121 GlyThrAspLeuValSerlleprohnglyprobsanvalThrValArgAlaenlleAla 140

```

Db      424 GGCACCGACCTGGTAGCATCCCAATGCGCCCAACGTCATCTGTCGTCACCAATTCCT 483
Qy      141 AAlaIeThrgIuSeArpLyBhePheIleAnGIySeArnTrpGluGlyIleGluGly 160
Db      484 GCCATCATGATGACAGACAAGTTCTTCATCAACGGCTCCAACTGGGAAGGACATCTCGGG 543
Qy      161 LeuAlaIyTAgIuIleAlaArpProApsApsSerLeuGluProPhePheApsSerLeu 180
Db      544 CTGGCTTATGCTGAGATTCGACGGCTGACGATCCCTGGAGCCCTTCTTGAATCTCTG 603
Qy      181 VAllyVsgInThrHlaValProApsLeuPheSerLeuGluLeuCYGlyValaGlyPhePro 200
Db      604 GTAAGGACAGACCCAGTTCCCACTCTTCTCTCCGACCTTGTGGGCTGGCTTCC 663
Qy      201 LeuAnGIuSeArpLyLeuAlaSeArValGIyGIySeArMeIleIleGIyGlyIleAap 220
Db      664 CTCAACGAGTGTGAAGTGTGGCTCTGTGCGAGGAGATGATCATTCAGAGATTCGAC 723
Qy      221 HisSerLeuYrThrGIySerLeuTrpYrThrProIleArGIuTrpYrThrGIu 240
Db      724 CACTCGCTGTACACAGGCGCTCTGTGTATACACCATCCGCGGAGGTGATTAATGAG 783
Qy      241 ValIleIleValArGIuValGIuIleAnGIyGIuApsLeuVseArPcyArGIuYr 260
Db      784 GTGATCATTTGGGGGTGGAGATCAATGACAGATCTGAAATGACCTGCAAGAGATTC 843
Qy      261 AsnYrAapLySerIleValApsSerGIyThrThrAsnLeuArGIuProIyArGIyVal 280
Db      844 AACTATGACAGAGCATTTCTGACAGTGGCACCAACCTTCGTTTGGCCAAAGAAAGTG 903
Qy      281 PheGIuAlaIaValIySerIleIyValaIaSerSerThrGIuLyPheProApsGIy 300
Db      904 TTGGAAGCTGCGAGTCAATCCATCAAGGACGCTCTCCACGAGAGTTCCTCATGTGT 963
Qy      301 PheTrpLeuGIyGIuGIuLeuValCYeTrpGIuAlaGIyThrThrProIyAsnIlePhe 320
Db      964 TTCTGGCTGAGAGAGAGCTGTGTGTGGCAAGGACCAACCCCTTGGACATTTTC 1023
Qy      321 ProValIleSerLeuYrLeuMeGIyGIuValIleThrAsnGIuSerPheArGIyleThrIle 340
Db      1024 CCAATCATTTCACTCACTAATGGGTAGGTACCAACAGATCCCTCCGATCACATC 1083
Qy      341 LeuProGIuGIuTrpLeuArGIuArGIuApsValaIleThrSerGIuApsPcyYr 360
Db      1084 CTTCGCGACCAATACCTGGCGGCGAGGAAAGTGTGCCACGTCCTCAAGACATGTTAC 1143
Qy      361 LysPheAlaIleSerGIuSerSerThrGIyThrValMeGIyValaValIleMeGIuGIy 380
Db      1144 AAGTTTGCATCTCAAGATTCACGGGCACTGTATGTGGAGCTGTATCATGGAGGGC 1203
Qy      381 PheYrValValPheApsArGIuArGIuArGIyleGIyPheAlaValSerAlaCYsHis 400
Db      1204 TTCTACGTTGTCTTGTATCGGGCCGAAAACAAATGGCTTGGCTCAGCGTTGCCAT 1263
Qy      401 ValHisApsGIuPheArGIuThrAlaValaGIuGIyProPheValIleLeuApsMeGIu 420
Db      1264 GTGCAAGATGAGTTCAAGCGGCAAGGCGGTGAAAGGCCCTTTGTCTTGGACATGGAA 1323
Qy      421 AapCYGlyYrAsnIleProGIuThrApsGIuSerThr 433
Db      1324 GACTGTGGCTACAAATTCACAGACAGATGAGTCAAC 1362

```

```

KW      muscular; antiparkinsonian; cerebroprotective; vasotropic; haemostatic;
KW      antidiabetic; antidiabetic; antidepressant; neurodegeneration; nerve damage;
KW      memory disorders; psychiatric disease; myopathy;
KW      mild cognitive impairment; Alzheimer's disease; as; gene; human.
OS      Homo sapiens.
FH      Key      Location/Qualifiers
FT      CDS      1..1527
FT      /*tag= a
XX      MO2004014843-A1.
XX      19-FEB-2004.
XX      07-AUG-2003; 2003WO-JP010045.
XX      09-AUG-2002; 2002JP-00233231.
XX      (TAKE ) TAKEDA CHEM IND LTD.
XX      Uchikawa O, Aso K, Koike T, Tani N, Hirai K;
XX      WPI; 2004-238691/22.
XX      P-PSDB; ADJ71858.
XX      New/known aryl carboxamide derivatives as inhibitors of aspartic acid
XX      protease and beta secretase for treating Alzheimer's disease,
XX      neurodegeneration, nerve damage, memory disorders, psychiatric disease,
XX      myopathy and cognitive impairment.
XX      Example 1; SEQ ID NO 6; 90pp; Japanese.
XX      The invention relates to novel N-Substituted aryl carboxamide compounds
XX      (I) and their salts. A compound of the invention has neuroprotective,
XX      neurotropic, neuroleptic, muscular-gen., antiparkinsonian,
XX      cerebroprotective, vasotropic, haemostatic, antidiabetic, antidiabetic, and
XX      antidepressant activity. The compounds are used to prevent and treat
XX      neurodegeneration, nerve damage, memory disorders, psychiatric disease,
XX      myopathy, mild cognitive impairment, or Alzheimer's disease, including
XX      Down's syndrome, senile dementia, Parkinson's disease, Creutzfeldt-Jacob
XX      disease, amyotrophic lateral sclerosis, diabetic neuropathy, Huntington's
XX      chorea, multiple sclerosis, cerebrovascular disorders, cerebral embolism,
XX      cerebral haemorrhage, cerebral arteriosclerosis, head injuries, spinal
XX      cord injuries, post-encephalitic disease, cerebral palsy, depression, panic
XX      disorder and schizophrenia. The present sequence is used in the
XX      exemplification of the invention.
SQ      Sequence 1527 BP; 317 A; 447 C; 438 G; 325 T; 0 U; 0 Other;

```

Alignment Scores:

Pred. No.:	2,09e-227	Length:	1527
Score:	2267.00	Matches:	429
Percent Similarity:	99.31%	Conservative:	1
Best Local Similarity:	98.08%	Mismatches:	3
Query Match:	99.08%	Indels:	0
DB:	12	Gaps:	0

US-10-726-967a-84 (1-433) x ADJ71857 (1-1527)

```

Qy      1 ThrgInHlaGIyTleArGIuProLeuArGIyLeuGIyValaProLeuGIyLeu 20
Db      64 ACCACAGCAGGATCCGCTGCTGAGAGTGTGTGGGGGGGCCCCCTGGGGCTG 123
Qy      21 GIuIleAenLeuGIuThrApsGIuGIuProGIuGIuProGIyArGIyArGIySerPheVal 40
Db      124 CGGCTGCCCGGGGAGACGAGAGAGCCGAGAGCCGCGGCGGAGGGGCGCTTGTG 183
Qy      41 GIuMetValaApsAenLeuArGIyLySerGIyGIuGIyTyrTyrValaGIuMetThrVal 60
Db      184 GAGATGTGTGACAACTGAGGCGCAAGTGTGGGCGAGGCTACTAGTGTGAGATGACCGTG 243
Qy      61 GlySerProProGIuThrLeuApsnIleLeuValAapThnGIySerSerApsnPheAlaVal 80

```



```

Db      124 CGGCTGCCCGGAGACCGACAGAGGCCGAGAGCCCGGAGGCGGACCTTGTG 183
      41 GImetValAspAsnLeuArgGlyLysSerGlyGlnGlyTyrTyrValGImetThrVal 60
Db      184 GAGATGCTGACAACTGAGGGGCAAGTGGGGGCAAGGCTTACGTGAGATGACCGTG 243
      61 GlySerProProGlnThrLeuAniLeuValAspThrGlySerSerAsnPheAlaVal 80
Db      244 GGCAGCCCCCGCAGACGCTCAACATCTGTGGATACAGGACGACGATTAATCTTGACAGTG 303
      81 GYAlaAlaProHisProPheLeuHisArgTyrTyrGlnArgGlnLeuSerSerThrTyr 100
Db      304 GGTGTGCCCCCGCAGCCCTTCTGTGATCGCTACTACAGAGGAGCTGTCCAGACATAC 363
      101 ArgAspLeuArgGlyValTyrValProTyrThrGlnGlyLysThrPGLugLugLeu 120
Db      364 CGGAGCTCCGAGAGGGGTGTATGTGCGCTTACACCCAGGGGCAAGTGGAGAGGAGCTG 423
      121 GYThrAspLeuValSerIleProHisGlyProAsnValThrValArgAlaAsnIleAla 140
Db      424 GGCACGCACTGTGAAGCATCCCGCCAGCCCAAGTCATGTGCGTGCCTCAACATTTGCT 483
      141 AlaIleThrGlnSerAspLysPhePheIleAsnGlySerAsnThrGlnGlyLysLeuGly 160
Db      484 GCCATCACTGAATCAAGACAAAGTTCTTCATCAACGGCTCCACCTGGAGAGGACATCTGGGG 543
      161 LeuAlaTyrAlaGlnIleAlaArgProAspAspSerLeuGlnProPhePheAspSerLeu 180
Db      544 CTGGCTTATGCTGAGATTTCCAGGCTTACGACTCCCTGGAGCTTTCTTATCTCTCTG 603
      181 ValLysGlnThrHisValProAsnLeuPheSerLeuGlnLeuGlyAlaGlyPhePro 200
Db      604 GTAAAGACAGACCAAGTCTCCCAACCTCTCTCCCTGCACTTGTGTGCTGCTGCTTCCCC 663
      201 LeuAsnGlnSerGlyValLeuAlaSerValGlyLysSerMetIleIleGlyLysLeuAsp 220
Db      664 CTCAACCAAGTCTGAAGTGTGCTGCTTGTGAGGAGGATATCATTTGAGATTCAGAC 723
      221 HisSerLeuTyrThrGlySerLeuTyrTyrThrProIleArgArgGlnTyrTyrGln 240
Db      724 CACTGCTTACACAGGAGCTCTCTGTATACCCATCCGCGGAGAGTGTATTTATGAG 783
      241 ValIleIleValArgValGlnIleAsnGlyGlnAspLeuLysMetAspCysLysGlnTyr 260
Db      784 GTGATCATTTGTGGGGTGGAGATCAATGACAGATCTGAATATGACGCAAGAGTAC 843
      261 AsnTyrAspLysSerIleValAspSerGlyThrThrAsnLeuArgLeuProLysVal 280
Db      844 AACTATGACAAAGACATTTGTGGACAGTGGACACCAACCTTGTGTGCTTGTGCTTGTGCT 903
      281 PheGlnAlaAlaValLysSerIleLysAlaIleSerSerThrGlnLysPheProAspGly 300
Db      904 TTGAAAGCTGCAATCAATTCATCAAGGAGCTCTCTTCCAGAGAGATTCCTGATGCT 963
      301 PheThrLeuGlyGlnGlnLeuValCysTyrGlnAlaGlyThrThrProTyrAsnIlePhe 320
Db      964 TTCTGCTAGAGAGAGAGCTGTGTGCTGTGCAAGGAGGACCAACCTTGTGAAATTTTC 1023
      321 ProValIleSerLeuTyrLeuMetGlyGlnValThrAsnGlnSerPheArgIleThrIle 340
Db      1024 CCAAGTATCTCACTTCACTTAAGGGGTGAGTTACCAACAGATCTTCCGATCACATC 1083
      341 LeuProGlnGlnTyrLeuArgProValGlnAspValAlaThrSerGlnAspAspCysTyr 360
Db      1084 CTTCGAGCAATACTGCGGCGAGTGGAGATGTGGCCAGTCCCAAGCAGATGTAC 1143
      361 LysPheAlaIleSerGlnSerSerThrGlyThrValMetGlyAlaValIleMetGlnGly 380
Db      1144 AACTTTGCATCTCAAGTCATCAAGGAGCATGTATATGAGACTGTATCATGAGAGGC 1203
      381 PheTyrValValPheAspArgAlaArgValArgIleGlyPheAlaValSerAlaCysHis 400

```

```

Db      1204 TTCTACGTGTTCTTGTGATCGGCGCCGAAACGAATGCTTGTGTCAGCGCTTGCCAT 1263
      401 ValHisAspGlnPheAspThrAlaAlaValGlnGlyProPheValThrLeuAspMetGln 420
Db      1264 GTGCACGATGAGTTTCAGAGCGCAGCGGTGGAAGGCTTTTGTGTCATCTTGACATGGAA 1323
      421 AspCysGlyTyrAsnIleProGlnThrAspGlnSerThr 433
Db      1324 GACTGTGCTACACATTTCCACAGACAGATGATGTAACC 1362
      RESULT 10
      ADH34044
      ID ADH34044 standard; cDNA; 1542 BP.
      AC ADH34044;
      DT 11-MAR-2004 (first entry)
      DE Human cDNA for His-tagged BACE (Beta-secretase).
      KW Beta-secretase; BACE; BACE-2; inhibitor; Alzheimer's disease;
      KW neuroprotective; nootropic; amyloid precursor protein;
      KW beta-amyloid peptide; cerebrovascular amyloidosis; ss; gene.
      OS Synthetic.
      OS Homo sapiens.
      PN US2003125257-A1.
      PD 03-JUL-2003.
      PF 18-DEC-2002; 2002US-00322684.
      PR 20-DEC-2001; 2001EP-00130282.
      XX
      PA (BROC/) BROCKHAUS M.
      PA (DOEB/) DOEBELI H.
      PA (GRUB/) GRUENINGER F.
      PA (HUGU/) HUGUENIN P.
      PA (KITA/) KITAS E A.
      PA (NELB/) NELBOECK-HOCHSTETTER P.
      PI Brockhaus M, Doebeli H, Grueninger F, Huguenin P, Kitas EA;
      PI Nelboeck-Hochstetter P;
      DR WPI; 2004-088811/09.
      PT New inhibitors of beta-secretase, useful for treating cerebrovascular
      PT amyloidosis, especially Alzheimer's disease, and for drug screening.
      PS Example 1; SEQ ID NO 1; 23pp; English.
      XX
      CC The invention relates to peptide beta-secretase (bs) inhibitors of
      CC genetic formula appearing as ADH34057. Also included are an assay for
      CC identifying inhibitors of bs, screening compounds for inhibition of bs
      CC activity, a kit for identifying a bs inhibitor and bs inhibitors
      CC identified using the kit. The bs used is isolated or recombinant, and
      CC purified, especially a full-length bs, specifically BACE or BACE-2. The
      CC inhibitors of beta-secretase (which is involved in degradation of amyloid
      CC precursor protein to beta-amyloid peptide) are used for treating patients
      CC with, or predisposed to, cerebrovascular amyloidosis, specifically
      CC Alzheimer's disease. They are also used to prepare tagged derivatives,
      CC useful in screening compounds for identifying other bs inhibitors and for
      CC radiolabeling or positron emission tomographic imaging. The present
      CC sequence encodes a His-tagged BACE protein used to produce BACE
      CC protein recombinantly.
      XX
      SQ Sequence 1542 BP; 322 A; 459 C; 433 G; 328 T; 0 U; 0 Other;
      Alignment Scores:
      Pred. No.: 2.12e-227 Length: 1542
      Score: 2267.00 Matches: 429
      Percent Similarity: 99.31% Conservative: 1

```



the vector and the method of producing Hu-Asp polypeptide, an isolated antibody that specifically binds to Hu-Asp polypeptides, identifying a cell that can be used to screen for inhibitors of beta secretase activity, novel isoforms of amyloid protein precursor (APP), where the last 2 carboxy terminus amino acids of that isoform are both lysine residues (e.g. those designated APP695-KK or carrying the Swedish mutation where Km at 595-596 is mutated to NL, designated e.g. APP65-Sw or APP695-Sw-KK, or a V to F mutation at 642, e.g. APP695-VF, all useful for assaying for beta secretase activity and screening for inhibitors of beta-secretase) and polynucleotides that encode the APP proteins. The method is useful for identifying agents that modulate the activity (amyloid precursor protein processing activity) of Asp2 aspartyl protease. Preferably, the method is useful for identifying agents that inhibit Asp2 aspartyl protease activity. The inhibitors of amyloid precursor protein processing, are useful for treating or preventing Alzheimer's disease. The present sequence encodes an aspartyl protease of the invention.

Sequence 2070 BP, 476 A, 582 C, 563 G, 449 T, 0 U, 0 Other;

## Alignment Scores:

Score: 3, 21e-227 Length: 2070  
Percent Similarity: 2267.00 Matches: 429  
Best Local Similarity: 99.31% Conservative: 1  
Query Match: 99.08% Mismatches: 3  
DB: 12 Indels: 0  
Gaps: 0

US-10-726-967a-84 (1-433) x ADJ94315 (1-2070)

```

QY 1 ThrGlnHISGLYTLeargleuProleuAArgSerGlyLeuGlyValAProleuGlyLeu 20
DB 64 ACCGACGACGCGATCGGCTGCGCCCTGCGACGCGGCTGGGGGCGCCCTGGGGGCTG 123
QY 21 GluIleAsnLeuGluThraBpGluGluProGluGluPProGluYArgArgGlySerPheVal 40
DB 124 CGGCTGCCCCGGGAGACCGACGAGAGAGCCGAGAGAGCCGCGCGAGGGGCGACTTTGTG 183
QY 41 GluMetValAspAsnLeuArgGlyLysSerGlyGlnGlyTyrTyrValGluMetThrVal 60
DB 184 GAGATGCTGACACACTGAGGGGCGAAGTGGGGCGAGGGCTACTGAGATGACCGTG 243
QY 61 GlySerProProGlnThraLeuAnlleuValAspThrGlySerSerAsnPheAlaVal 80
DB 244 GCGAGGCCCCCGACGAGCTCAACATCTGCTGATGATACAGCGACGATTAATTGACATG 303
QY 81 GlyAlaAlaProHisAspPheLeuHisArgTyrTyrGlnArgGlnLeuSerSerThrTyr 100
DB 304 GGTGCTGCCCCCGACCCCTTCTGCTGATCGTACTACAGAGCGAGCTGTCCAGCACATAC 363
QY 101 ArgAspLeuArgGlyValTyrValProTyrThrGlnGlyLysTyrPProGluGluLeu 120
DB 364 CGGAGCTCCGAGAGGCTGTATGTGCTTACACCCAGGCGCAAGTGGAGAGGGAGCTG 423
QY 121 GlyThrAspLeuValSerIleProHisGlyProAsnValThrValArgAlaAsnIleAla 140
DB 424 GGGACCGACCTGTGTATGATCCCGCATGGCCCAAGCTCATCTGTGCTGCGCAATTTGCT 483
QY 141 AlaIleThrGluSerAspLysPhePheIleAsnGlySerAsnTyrPProGluGlyIleLeuGly 160
DB 484 GCCATCACTGAAATCAACAGATTTCTTCATCAACGAGCTCCAACTGGAGAGCATCTGGGG 543
QY 161 LeuAlaTyrAlaGluIleAlaArgProAspAspSerLeuGluProPhePheAspSerLeu 180
DB 544 CTGGCTTATGCTAGATTTGCGACGCTGACGACTCTCTGAGGCTTTCTTTGACTCTCTG 603
QY 181 ValIleGlnThrHisValProAsnLeuPheSerLeuGlnLeuGlyGlyAlaGlyPhePro 200
DB 604 GTAACGACAGACCGACGTTCCCAACTCTTCTCCCTGAGCTTGTGTGTGCTGCTTCC 663
QY 201 LeuAsnGlnSerGluValIleuAlaSerValGlyGlySerSerIleIleGlyGlyIleAsp 220
DB 664 CTCACCAAGCTGTAAGTGTGCTGCTGTGTGAGAGGAGCATGATTCATGGAGGATTCGAC 723

```

```

QY 221 HisSerLeuTyrThrGlnSerLeuTyrThrProIleArgArgGluTyrTyrGlu 240
DB 724 CACTGGCTGACAGGAGAGTCTCTGTATACACCATCCGGGAGGTGATTTATAG 783
QY 241 ValIleIleValArgValGluIleAsnGlyGlnAspLeuLysMetAspCysLeuGlyTyr 260
DB 784 GTCATCATTTGGCGGGTGGATCATATGACAGGATTCGAAATGACTGCACAGGAGTAC 843
QY 261 AsnTyrAspLysSerIleValAspSerGlyThrThrAsnLeuArgLeuProLysVal 280
DB 844 AACTATGACAAAGAGATTTGTGACAGTGGCACACCAACCTTCCTTCCCAAGAAATG 903
QY 281 PheGluAlaAlaValIleSerIleValAlaIleSerSerThrGluLysPheProAspGly 300
DB 904 TTGGAAGCTGACGTAATTCATCAAGGAGCGCTCTTCCAGGAGAAATTCCTGATGCT 963
QY 301 PheTyrLeuGlyGluGlnLeuValCysTyrPProGlnAlaGlyThrThrProTyrAsnIlePhe 320
DB 964 TTCTGGCTAGAGAGACACTGTGTGCTGGCAGAGGACACACCCCTTGGAACTTTTC 1023
QY 321 ProValIleSerLeuTyrLeuMetGlyGluValThrAsnGlnSerPheArgIleThrIle 340
DB 1024 CCAATCATCTCACTTACTTAAATGGGTGAGGTACCAACAGTCTTCCGATCACCATC 1083
QY 341 LeuProGlnGlnTyrLeuAspProValGluAspValAlaThrSerGlnAspAspCysTyr 360
DB 1084 CTTCCGCGACATTAATCTCGGCGAGTGAAGATGGGCCAGCTCCCAAGACGACTTTTC 1143
QY 361 LysPheAlaIleSerGlnSerSerThrGlyThrValMetGlyAlaValIleMetGlyGly 380
DB 1144 AAGTTTGCATTTCACTATTCACAGGCACTGTATGGAGCTGTATATGAGAGGCG 1203
QY 381 PheTyrValIlePheAspArgIleArgLysArgIleGlyPheAlaValSerAlaCysHis 400
DB 1204 TTCTACGTTGCTTGTGATCGGGCCGAAACGAATTTGCTTGTGCTGACGCCCTTGCCAT 1263
QY 401 ValHisAspGluPheArgThrAlaAlaGluGlyPProPheValThrLeuAspMetGlu 420
DB 1264 GTGACAGATGATTCAGAGCGGACGGTGGAGGCCCTTTGTGACCTTGGACATGGAA 1323
QY 421 AspCysGlyTyrAsnIleProGlnThrAspGluSerThr 433
DB 1324 GACTGTGCTTACATTCACAGACAGATGATGTAACC 1362

```

RESULT 12  
AD050411  
ID AD050411 standard; cDNA; 2070 BP.  
XX  
AC AD050411;  
XX  
DT 29-JUL-2004 (first entry)  
XX  
DE Human aspartyl protease (Asp)-2(a) cDNA.  
XX  
KW Aspartyl protease; Asp; beta secretase; amyloid precursor protein; APP;  
KW Alzheimer's disease; gene therapy; human; gene; chromosome 11q23.3-24.1;  
KW ss.  
XX  
OS Homo sapiens.  
XX  
XX  
XX Key Location/Qualifiers  
XX FT 1..1506  
XX FT /\*tag= b  
XX FT /product= "Human Asp-2 protein"  
XX FT sig\_peptide 1..63  
XX FT /\*tag= a  
XX FT mat\_peptide 64..1503  
XX FT /\*tag= C  
XX FT /product= "Human mature Asp-2 protein"  
XX  
XX US6737510-B1.  
XX



PD 18-MAY-2004.

PF 12-APR-2000; 2000US-00548373.

XX 24-SEP-1998; 98US-0101594P.

XX 23-SEP-1999; 99US-00404133.

PR 23-SEP-1999; 99US-0155493P.

PR 23-SEP-1999; 99WO-US020881.

PR 13-OCT-1999; 99US-00416501.

XX (PHAA ) PHARMACIA & UPJOHN CO.

PA Gurney ME, Blenkowiaki MJ, Heinrikson RL, Parodi LA, Yan R;

XX MPI: 2004-387112/36.

XX P-PSDB; ADO50412.

DR

XX

PT New Asp2 aspartyl protease protein comprising tripeptides DTG and DSG

PT involved in processing amyloid precursor protein into amyloid beta,

PT useful in preparing a composition for treating or preventing Alzheimer's

PT disease.

XX

PS Example 2; SEQ ID NO 3; 108bp; English.

XX

CC The invention relates to a method for identifying an agent that decreases

CC the protease activity of the aspartyl protease (Asp) polypeptide. It also

CC provides enzyme and enzymatic procedures for cleaving the beta secretase

CC cleavage site of the amyloid precursor protein (App). The invention is

CC useful in preparing a composition for treating or preventing Alzheimer's

CC disease. It is also useful in gene therapy. The present sequence is human

CC Asp-2 CDNA. Human Asp-2 gene is located at chromosome 11q23.3-24.1. This

CC sequence is used to illustrate the method of the invention.

XX

SQ Sequence 2070 BP; 476 A; 582 C; 563 G; 449 T; 0 U; 0 Other;

Alignment Scores:

Prod. No.: 3,21e-227 Length: 2070

Score: 2267.00 Matches: 429

Percent Similarity: 99.31% Conservative: 1

Best Local Similarity: 99.08% Mismatches: 3

Query Match: 99.08% Indels: 0

DB: 12 Gaps: 0

US-10-726-967A-84 (1-433) x ADO50411 (1-2070)

QY 1 ThrGlnHtGtGlyLLeHrGleuProLeuArgSerGlyLeuGlyGlyAlaProLeuGlyLeu 20

Db 64 ACCGAGCAGCGGCAATCGGCTGCCCTGCGACGGCGCGGGGGGGCCCCCTTGGGGCTG 123

QY 21 GluIleAenLeuGluThrAspGluGluProGluGluProGluArgArgGlySerPheVal 40

Db 124 CGGCTGCCCGGGAGAGCCGACCAAGAGCCCGAGAGCCCGCGGAGGGGCAAGCTTTGTG 183

QY 41 GluMetValAspAsnLeuArgGlySerGlyGlnGlyTyrTyrValGluMetThrVal 60

Db 184 GAGATGATGACAACTCGAGGGGCAAGTCGGGGGCGGGCTACTCGTGAGATGACCGTG 243

QY 61 GlySerProProGluThrLeuAsnIleLeuValAspThrGlySerSerAsnPheAlaVal 80

Db 244 GGGAGCCCCCGGAGCGCTCAACATCTGGGGATACAGGACAGCACTTTGCAGTG 303

QY 81 GlyAlaAlaProHisProPheLeuHisArgTyrTyrGlnArgGlnLeuSerSerThrTyr 100

Db 304 GGTGTGCCCCCGACCCCTTCTGTGATGGCTACTACAGAGGAGCTGTCCAGACACTAC 363

QY 101 ArgAspLeuArgGlyGlyValTyrValProTyrThrGlnGlyTyrTPGJuglyGluLeu 120

Db 364 CGGGAAGCTCCGGAAGGTGTGTATGTGGCTTACACCCGAGGCAAGTGGAAAGGGAGCTG 423

QY 121 GlyThrAspLeuValSerIleProHisGlyProAsnValThrValArgAlaAsnIleAla 140

Db 424 GGGACGAGCTGTGATAGCATCCCCCAAGGCCCAAGCATGAGCTGTGGCTGCAACTTGGCT 483

QY	141	AlaIleThrGlnSerAspIysPhePheIleamGlySerAsnTrpGluIylIleuGly	160
Db	484	GCCATCACTGAATCAACAAGATTCTTCATCAACGGCTCCAACTGGGAAGGCATCTCGGGG	543
QY	161	LeuAlaTYAlaGluIylIeAlaArgProAspAspSerLeuGluProPhePheAspSerLeu	180
Db	544	CTGGCTCATTCGAGATTGCGAGGCTGACGACTCCCTGGAGGCTTTCTTTGACTCTCG	603
QY	181	ValIysGlnThrHisValProAsnLeuPheSerLeuGlnLeuCyGlyAlaGlyPhePro	200
Db	604	GTAACCAACACCACGCTCCCAACTCTTCTCCCTGACACTTTGGTGCTGCTGCTTCCC	663
QY	201	LeuAsnGlnSerGluValIleuAlaSerValGlyGlySerMetIleIleGlyGlyIleAsp	220
Db	664	CTCAACCACTCGAAGTGTGGCTGCTGCTGCTGCGAGGAGCATGATCATTTGAGATCGAC	723
QY	221	HisSerLeuTYThrGlySerLeuTrpTYThrProIleArgArgGluTrpTYTrpGlu	240
Db	724	CACCTCGCTGTACAACAGGACGTCTCTGGTATACACCATCGGCGGAGGTGATATATAG	783
QY	241	ValIleIleValArgValGluIylIeamGlyIylAsnLeuLysMetAspCysIleGlyTrp	260
Db	784	GTCATCATTTGTGGGGTGGAGATCAATGACAGAGATCTGAATAATGACATGCAAGAGTAC	843
QY	261	AsnTYAspLysSerIleValAspSerGlyTYThrAsnLeuArgLeuProLysIleVal	280
Db	844	AACATATGACAAAGACATTGTGGACAGTGGCACCAACCACTTCGTTTGGTCCCAAGAAAGTG	903
QY	281	PheGluAlaIleValLysSerIleIylAlaIleAspSerThrGluLysPheProAspGly	300
Db	904	TTTGAAGCTGACGTCAAAATCCATCAAGGACGCTCTCCACGGAAGATGCCATAGGT	963
QY	301	PheTrpLeuGlyGluGlnLeuValCysTrpGlnAlaGlyTYThrProTrpAsnIlePhe	320
Db	964	TTCTGGCTAGAGAGACAGCTGGTGTGCTGGCAAGCACCAACCCCTTGGAACATTTTC	1022
QY	321	ProValIleSerLeuTYLeuMetGlyGluValIleAsnGlnSerPheArgIleThrIle	340
Db	1024	CCAGCATCACTCACTTACCTATGGGTGAGGTTACCAACGAGTCTTCGCGATCACCATC	1083
QY	341	LeuProGlnGlnTYLeuArgProValGluAspValAlaThrSerGlnAspAspCysTYr	360
Db	1084	CTTCGCGACGAACTACTCGGCGCAATGGGAAGTGTGGCACGCTCCCAAGACGACTGTAC	1143
QY	361	LysPheAlaIleSerGlnSerSerThrGlyTYThrValMetGlyAlaValIleMetGluGly	380
Db	1144	AAGTTTGCATCTCAAGTCATCCACGGGCACTGTTATGGGAGCTTATCATGAGGGC	1203
QY	381	PheTYrValIleValPheAspArgAlaArgIylAsnArgIleGlyPheAlaValIleSerIleCysHis	400
Db	1204	TTCTACGTGTCTTGTGATCGGGCCCGAAACGAATGTGCTTGTGCTGACGAGGCTTGCAT	1263
QY	401	ValHisAspGluPheArgTYrAlaIleValGluGlyProPheValIleLeuAspMetGlu	420
Db	1264	GTCACGATGAGTTTCAGGACGGGACGGGTGGAAGGCCCTTTTGTGCACCTTGACATGGA	1323
QY	421	AspCysGlyTYrAsnIleProGlnThrAspGlnSerThr	433
Db	1324	GACTGTGGCTCAACATTCACAGACGATGATCAACCC	1362
RESULT 13			
ADR75324			
ID	ADR75324	standard; cDNA, 2070 BP.	
XX	ADR75324;		
XX	18-NOV-2004	(first entry)	
XX	Human aspartyl protease (Asp-2(a))	CDNA.	
XX	Aspartyl protease; Asp; amyloid precursor protein; APP; amyloid beta;		
XX	chromosome identification; Alzheimer's disease; human; gene; ss.		





QY 401 ValHisAspGluPheArgThrAlaAlaValGluGlyProPheValThrLeuAspMetGlu 420  
|||  
Db 1264 GTGCACGATGAGTTCAGACCGGACCGGTGAAGGCCCTTTGTGCACCTGGACATGGAA 1323  
QY 421 AApCySGlyTYrAsnIleProGlnThrAspGluSerThr 433  
|||  
Db 1324 GACTGTGCTACACATTCACACGACGATGAGTCAACC 1362  
RESULT 14  
AAAS9551  
ID AAAS9551 standard; DNA; 2348 BP.  
XX  
AC AAAS9551;  
XX  
DT 14-NOV-2000 (first entry)  
XX  
DE DNA encoding a human beta-secretase enzyme.  
XX  
KM Beta-secretase; beta-amyloid precursor protein; beta-amyloid peptide;  
KM amyloid plaque component; Alzheimer's disease; amyloidogenic disease;  
KM inhibitor; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 306..1811  
FT /\*tag= a  
FT /product= "beta-secretase"  
XX  
PN MO200047618-A2.  
XX  
PD 17-AUG-2000.  
XX  
PF 10-FEB-2000; 2000WO-US003819.  
XX  
PR 10-FEB-1999; 99US-0119571P.  
PR 15-JUN-1999; 99US-0139172P.  
XX  
PA (ELAN-) ELAN PHARM INC.  
XX  
PI Anderson JP, Baal G, Doane MT, Frigon N, John V, Power M;  
PI Simha S, Tatsuno G, Tung J, Wang S, Mcconlogue L;  
XX  
DR WPI: 2000-533011/48.  
DR P-PSDB; AAB07896.  
XX  
PT Purified beta-secretase protein used in assays to discover inhibitors  
PT which can be used for the treatment of amyloidogenic diseases e.g.  
PT Alzheimer's disease.  
XX  
PS Disclosure; Fig 1B; 121pp; English.  
XX  
CC The specification describes a beta-secretase enzyme. The enzyme cleaves  
CC beta-amyloid precursor protein to produce beta-amyloid peptide. This  
CC enzyme is therefore implicated in the production of amyloid plaque  
CC components which accumulate in the brains of individuals afflicted with  
CC Alzheimer's disease. Inhibitors of beta-secretase are administered to a  
CC mammalian subject e.g. with Alzheimer's disease or Alzheimer's disease-  
CC like pathology to test if they maintain or improve cognitive ability or  
CC reduce the plaque burden. The compounds are used for the treatment of  
CC amyloidogenic diseases e.g. Alzheimer's disease. The present sequence  
CC encodes a human beta-secretase enzyme  
XX  
SQ Sequence 2348 BP; 489 A; 713 C; 661 G; 484 T; 0 U; 1 Other;

Alignment Scores:  
Pred. No.: 3.83e-227 Length: 2348  
Score: 2267.00 Matches: 429  
Percent Similarity: 99.31% Conservative: 1  
Best Local Similarity: 99.08% Mismatches: 3  
Query Match: 99.08% Indels: 0  
DB: 3 Gaps: 0

US-10-726-967a-84 (1-433) x AAAS9551 (1-2348)  
QY 1 ThrGlnHisGlyIleArgLeuProLeuArgSerGlyLeuGlyGlyAlaProLeuGlyLeu 20  
|||  
Db 369 ACCCAGACCGGCATCCCGCTGCCCTCCGCGACGGCCCTGGGGGGGCCCCCTGGGGCTG 428  
QY 21 GluIleAsnLeuGluThrAspGluGluProGluGluProGlyArgArgGlySerPheVal 40  
|||  
Db 429 CGGCTGCCCGGAGACCGACGAGAGGCCGAGAGGCCCGGAGGGGACCTTTGTG 488  
QY 41 GluMetValAspAsnLeuArgGlyLeuSerGlyGlnGlyTYrTYrValGluMetThrVal 60  
|||  
Db 489 GAGATGGTGAACAACCTGAGGGGCAAGTCCGGGCGAGGGCTACTACGTGAGATGACCGTG 548  
QY 61 GlySerProProGlnThrLeuAsnIleLeuValAspThrGlySerSerAsnPheAlaVal 80  
|||  
Db 549 GGCAGCCCCCGGAGAGCTCCACATCTGTGTGATACAGGCAAGTAATTGGCAGTG 608  
QY 81 GluAlaAlaProHisProPheLeuHisArgTYrTYrGlnArgGlnLeuSerSerThrTYr 100  
|||  
Db 609 GGTCTGCCCCCGACCCCTTCCTGCACTGCTACTACAGAGGACGCTGTCAGACATAC 668  
QY 121 GlyThrAspLeuValSerIleProHisGlyProAsnValThrValAlaGlnIleAla 140  
|||  
Db 729 GGCACCCACCTGGTACATCCCTCCATGGCCCCCAACCTCACTGTCGTGCCAACATTTGCT 788  
QY 141 AlaIleThrGluSerAspIlePhePheIleAsnGlySerAsnTPGluGlyIleLeuGly 160  
|||  
Db 789 GCCATCACTGATACAGACAACTTTTCATCAACGGCTCCAACTGGGAAGGATCTTGGG 848  
QY 161 LeuAlaTYrAlaGluIleAlaArgProAspAspSerLeuGluProPheAspSerLeu 180  
|||  
Db 849 CTGGCTATGTGTGAGATTGGCAGGCTGACGATCCCTGGAGGCTTTCTTGACTCTCTG 908  
QY 181 ValIleGlnThrHisValProAsnLeuPheSerLeuGlnLeuCyGlyAlaGlyPhePro 200  
|||  
Db 909 GTAAAGACAGACCCACGTTCCACACTCTCTCCCTGACGTTGTGGTGGCTTCC 968  
QY 201 LeuAsnGlnSerGluValIleAsnIleSerValGlyGlySerMetIleIleGlyIleAsp 220  
|||  
Db 969 CTCAACCAAGTCTGAAGTGTGGCTCTGTGCGAGGAGCATGATCATTTGAGAGTATGAC 1028  
QY 221 HisSerLeuTYrThrGlySerLeuTPYrThrProIleArgArgGluTPYrTYrGlu 240  
|||  
Db 1029 CACTCGCTATACACAGGCAAGCTCTGGTATACACCATCCGGCGGAGTGTATTATGAG 1088  
QY 241 ValIleIleValArgValGluIleAsnGlyGlnAspLeuIleMetAspCyGlyGluTYr 260  
|||  
Db 1089 GTGATCATTTGGCCGGTGGAGATCAATGACAGATCTGAAATGACCTCCAAAGACTAC 1148  
QY 261 AsnTYrAspIleSerIleValAspSerGlyThrThrAsnLeuArgLeuProIleGlyVal 280  
|||  
Db 1149 AACATGACAAAGACATTTGAGACAGTGGCACCAACCACTTCCTTTGCCCAAGAAATG 1208  
QY 281 PheGluAlaAlaValIleSerIleIleValAlaIleSerSerThrGluIlePheProAspGly 300  
|||  
Db 1209 TTTAAGCTGACAGCAATATCATCAAGGACGCTCCCTCCAGGAGAAATTCCTGATGCT 1268  
QY 301 PheTYrLeuGlyGluGlnLeuValCyeTYrGlnAlaGlyTYrThrProTPAsnIlePhe 320  
|||  
Db 1269 TTTCTGGCTAGAGAGCAGCTGTGTGTGCTGGCAGACAGCAACCCCTTGGAACATTTTC 1328  
QY 321 ProValIleSerLeuTYrLeuMetGlyGluValThrAsnGlnSerPheArgIleThrIle 340  
|||  
Db 1329 CCAATCATCTACCTACTATATGAGGTGAGGTTCACCAACGATCTTCGGATACCATTC 1388  
QY 341 LeuProGlnGlnTYrLeuArgProValGluAspValaIleThrSerGlnAspAspCyTYr 360  
|||

Db 1389 CTTCCGACGACATACCTGGCGCAGTGCAGAAATGTGGCCACGATCCAGACGACTGTAC 1448  
Qy 361 TysPheAlaIleSerGlnSerSerThrGlyThrValMetGlyAlaValIleMetGluGly 380  
Db 1449 AAGTTGGCATCTCAACAGTCATCCACGCGCAGCTTATGTGGAGCTGTATCATGAGGCGC 1508  
Qy 381 PheTyrValValPheAspArgAlaArgIleGlyPheAlaValSerAlaCysHis 400  
Db 1509 TTCTACGTTGTCTTGTATCGGCGCCGAAACGAAATGGCTTGTGCTGACGCTTGCCAT 1568  
Qy 401 ValHisAspGluPheArgThrAlaAlaValGluGlyProPheValThrLeuAspMetGlu 420  
Db 1569 GTGCACGATGATTCAGACGCGCAGCGGTGGAAGGCCCTTTGTACCTTGGACATGAA 1628  
Qy 421 AspCysGlyTyrAsnIleProGlnThrAspGluSerThr 433  
Db 1629 GACTGTGGCTACACATTCACAGACGATGAGTCAACC 1667  
RESULT 15  
ID AAS82237/c  
XX AAS82237 standard; cDNA; 2907 BP.  
AC AAS82237;  
XX  
DT 13-FEB-2002 (first entry)  
XX  
DE DNA encoding novel human diagnostic protein #18041.  
XX  
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX  
OS Homo sapiens.  
XX  
PN MO200175067-A2.  
XX  
PD 11-OCT-2001.  
XX  
PF 30-MAR-2001; 2001WO-US008631.  
XX  
PR 31-MAR-2000; 2000US-00540217.  
XX  
PR 23-AUG-2000; 2000US-00649167.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Dmanac RT; Liu C; Tang YT;  
XX  
DR WPI; 2001-639362/73.  
XX  
DR P-PsDB; ABG18050.  
XX  
PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.  
XX  
PS Claim 1; SEQ ID NO 18041; 103pp; English.  
XX  
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantifying a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic

CC coding sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 2907 BP; 623 A; 801 C; 811 G; 672 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 5,18e-227 Length: 2907  
Score: 2267.00 Matches: 429  
Percent Similarity: 99.31% Conservative: 1  
Best Local Similarity: 99.08% Mismatches: 3  
Query Match: 99.08% Indels: 0  
DB: 5 Gaps: 0  
US-10-726-967A-84 (1-433) x AAS82237 (1-2907)  
Qy 1 ThrGlnHisGlyIleArgLeuProLeuArgSerGlyLeuGlyAlaProLeuGlyLeu 20  
Db 2844 ACCCAGACGCGGATCCGGCTGCCCCGCGACGCGGCTGGGGGGGCGCCCCCTGGGGCGTG 2785  
Qy 21 GluIleAsnLeuGluThrAspGluGluProGluGluProGlyArgArgGlySerPheVal 40  
Db 2784 CGGCTGCCCGGGAGACCGACGAGAGCGCCGAGGCGCCGGCGGAGGGCGAGCTTTGTG 2725  
Qy 41 GluMetValAspAsnLeuArgGlyIleSerGlyGlnGlyTyrTyrValGluMetThrVal 60  
Db 2724 GAGATGTGTGACACCTGAGGGGAGATGGGGGCGGCGCTACTACGTGAGATGACCGTG 2665  
Qy 61 GlySerProProGlnThrLeuAsnIleLeuValAspThrGlySerSerAsnPheAlaVal 80  
Db 2664 GGCAGCCCCCGGACACCTCAACATCCTGGTGGATACAGGACGACGATACCTTGCAGTG 2605  
Qy 81 GlyAlaAlaProHisProPheLeuHisArgTyrTyrGlnArgGlnLeuSerSerThrTyr 100  
Db 2604 GGTGCTGCCCGCCACCCCTTCTGATGCTACTACAGAGGACGACCTGTCCAGCAATAC 2545  
Qy 101 ArgAspLeuArgGlyGlyValTyrValProTyrThrGlnGlyIleTyrGluLeu 120  
Db 2544 CGGGAACCTCCGAAAGGTGTGTATGTGCTTACACCCAGGGGAGATGGAGGGAGCTG 2485  
Qy 121 GlyThrAspLeuValSerIleProHisGlyProAsnValThrValArgAlaAsnIleAla 140  
Db 2484 GGCACCGACCTGTAAAGATCCCCATGGCCCAAGTCATCTGTGGTCCAAATGTGCT 2425  
Qy 141 AlaIleThrGluSerAspIlePhePheIleAsnGlySerAsnTyrGluIleGluGly 160  
Db 2424 GCCATCAGTGAATCAGACAAAGTTCTTCAATCAACGCTCCAACTGGGAGGCAATCTGGGG 2365  
Qy 161 LeuAlaTyrAlaGluIleAlaArgProAspAspSerLeuGluProPheAspSerLeu 180  
Db 2364 CTGGCTTATGTGAGATTTGCCAGGCTGACGACTCCCTGGAGCTTTCTTTATCTCTGTG 2305  
Qy 181 ValIleValThrHisValProAsnLeuPheSerLeuGlnLeuCysGlyAlaGlyPhePro 200  
Db 2304 GTAAAGCAGACCCAGCTTCCAACTCTTCTCCCTGACGCTTTGTGTCTGGCTTCCCC 2245  
Qy 201 LeuAsnGlnSerGluValLeuAlaSerValGlyGlySerMetIleIleGlyGlyIleAsp 220  
Db 2244 CTCACACATCTGAAGTGTGCTGTGCTGTGCGAGGGAGATCATTCATGAGATATCGAC 2185  
Qy 221 HisSerLeuTyrThrGlySerLeuTyrTyrThrProIleArgArgGluTyrTyrGlu 240  
Db 2184 CACTCGCTGTACACAGGCGTCTCGTGTATACCCATCCGCGGAGGTGATTAATGAG 2125  
Qy 241 ValIleIleValArgValGluIleAsnGlyIleAspLeuIleMetAspCysIleGlyTyr 260  
Db 2124 GTGATATTGTGGGTGTGAGATCAATGACAGATCTAATAATGACTGCAAGAGATAC 2065  
Qy 261 AsnTyrAspIleSerIleValAlaSerGlyTyrThrAsnLeuArgLeuProIleGlyVal 280  
Db 2064 AACTATGACAAAGCATTTGTGACAGTGCACACCAACCTTCGTTGTCACAAAGATG 2005

```
QY 281 PheGluAlaValLysSerLLeValAlaSerSerThrGluLysPheProAspGly 300
    |||
Db 2004 TTGAGCTGCGAATCCATCAAGGAGCCTCTCCACGAGAGATTCCCTGATGAT 1945
QY 301 PheTrpLeuGlyGluGlnLeuValCysTrpGlnAlaGlyThrThrProTrpAsnLLePhe 320
    |||
Db 1944 TTCTGGCTAGAGAGACAGCTGCTGTGCTGGCAAGCAGGACACCCCTTGGAAACATTTTC 1885
QY 321 ProValLLeSerLeuTrpLeuMetGlyGluValThrAsnGlnSerPheArgLLeThrLLe 340
    |||
Db 1884 CCAGTCATCTCACTCTACTTAATGGGTGAGTTACCAACAGTCCTCTCCGATCACATC 1825
QY 341 LeuProGlnGlnTrpLeuArgProValGluAspValAlaThrSerGlnAspAspCysTyr 360
    |||
Db 1824 CTTCCGACGCAATACCTGGGCGCAGTGGAGAGTGGCCACGTCCTCCAGACGACTGTTC 1765
QY 361 LysPheAlaLLeSerGlnSerSerThrGlyThrValMetGlyAlaValLLeMetGluGly 380
    |||
Db 1764 AAGTTGCCATCTCAACAGTCATCCACGGGCACTGTTATGGAGCTGTATCATGGAGGGC 1705
QY 381 PheTrpValValPheAspArgAlaArgLysArgLLeGlyPheAlaValSerLLeCysHis 400
    |||
Db 1704 TTCTACGTTGCTTGTGATCGGGCCGAAACGAATTGGCTTGTCTGACGCGCTTGCCAT 1645
QY 401 ValHisAspGluPheArgThrAlaAlaValGluGlyProPheValThrLeuAspMetGlu 420
    |||
Db 1644 GTGCGAGATGAGTTCAGAGCGGCAAGGAGGCGCTTTGTGACCTTGGAACATGGAA 1585
QY 421 AspCysGlyTyrAsnLLeProGlnThrAspGlnSerThr 433
    |||
Db 1584 GACTGTGGCTACAACTTCACAGACAGATGATCAACC 1546
```

Search completed: August 22, 2005, 01:11:01  
Job time : 1006 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 22, 2005, 00:24:09 ; Search time 8190 Seconds  
(without alignments)  
2561.796 Million cell updates/sec

Title: US-10-726-967A-84

Perfect score: 2268  
Sequence: 1 TQHGIRPLRSGLGAPLGL.....FTVLDMECGNINPQDEST 433

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=x1p  
-Q=/cgn2.1/USPTO.spool.p/US10726967/rnat.19082005.201240.29623/app.query.fasta.1.583  
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=45  
-UNITS=bits -STRM=1 -END=1 -MATRIX=blcosum62 -TRANS=human0.cdt -LIST=45  
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=PCO -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10726967@cgn.1.1.5600@rnat.19082005.201240.29623 -NCPU=6 -ICPU=3  
-NO WMAP -LARGEBUFFER -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOC  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:.\*  
1: gb\_ba:.\*  
2: gb\_hgt:.\*  
3: gb\_in:.\*  
4: gb\_cm:.\*  
5: gb\_ov:.\*  
6: gb\_pat:.\*  
7: gb\_ph:.\*  
8: gb\_pl:.\*  
9: gb\_pr:.\*  
10: gb\_ro:.\*  
11: gb\_sls:.\*  
12: gb\_sy:.\*  
13: gb\_un:.\*  
14: gb\_vl:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2267	99.1	1503	6 AR404163	AR404163 Sequence
2	2267	99.1	1506	6 AX700448	AX700448 Sequence
3	2267	99.1	1506	9 AF204943	AF204943 Homo sapi
4	2267	99.1	1527	6 BD103723	BD103723 Beta Secr

5	2267	99.1	2070	6 AR224093	AR224093 Sequence
6	2267	99.1	2070	6 AR269224	AR269224 Sequence
7	2267	99.1	2070	6 AR478778	AR478778 Sequence
8	2267	99.1	2070	6 AR487345	AR487345 Sequence
9	2267	99.1	2070	6 AR531984	AR531984 Sequence
10	2267	99.1	2070	6 AR540886	AR540886 Sequence
11	2267	99.1	2070	6 AR560095	AR560095 Sequence
12	2267	99.1	2070	6 AX700447	AX700447 Sequence
13	2267	99.1	2070	6 AF200343	AF200343 Homo sapi
14	2267	99.1	2348	6 AR404203	AR404203 Sequence
15	2267	99.1	2348	6 AR404204	AR404204 Sequence
16	2267	99.1	2526	6 AX700446	AX700446 Sequence
17	2267	99.1	2526	6 AF190725	AF190725 Homo sapi
18	2267	99.1	3252	6 AR305033	AR305033 Sequence
19	2267	99.1	3252	6 AX062111	AX062111 Sequence
20	2267	99.1	3252	6 AX063201	AX063201 Sequence
21	2267	99.1	3252	6 AX472368	AX472368 Sequence
22	2267	99.1	3252	6 AX700453	AX700453 Sequence
23	2267	99.1	3252	9 AF200193	AF200193 Homo sapi
24	2267	99.1	3893	9 BC065492	BC065492 Homo sapi
25	2267	99.1	5832	6 CO824591	CO824591 Sequence
26	2267	99.1	5842	6 CO726974	CO726974 Sequence
27	2267	99.1	5878	9 AF201468	AF201468 Homo sapi
28	2267	99.1	16080	6 AR404205	AR404205 Sequence
29	2267	98.9	1368	6 CO772936	CO772936 Sequence
30	2262	98.9	2070	6 BD235886	BD235886 Sequence
31	2262	98.9	2070	6 AX105385	AX105385 Sequence
32	2262	98.9	2070	6 AX378015	AX378015 Sequence
33	2262	98.9	2070	6 AX573823	AX573823 Sequence
34	2261	98.8	1506	6 AX823518	AX823518 Sequence
35	2261	98.8	2174	9 BC036084	BC036084 Homo sapi
36	2261	98.8	2541	6 AR178469	AR178469 Sequence
37	2261	98.8	2541	6 ES0816	ES0816 Aspartate p
38	2261	98.8	2541	6 AX002655	AX002655 Sequence
39	2261	98.8	2541	6 AX700452	AX700452 Sequence
40	2259	98.7	1368	6 CO772942	CO772942 Sequence
41	2257	98.6	1302	6 BD235897	BD235897 Alzheim
42	2257	98.6	1302	6 AR224103	AR224103 Sequence
43	2257	98.6	1302	6 AR269234	AR269234 Sequence
44	2257	98.6	1302	6 AR478789	AR478789 Sequence
45	2257	98.6	1302	6 AR487355	AR487355 Sequence

## ALIGNMENTS

RESULT 1  
LOCUS AR404163 1503 bp DNA linear PAT 18-DEC-2003  
DEFINITION Sequence 1 from patent US 6627739.  
ACCESSION AR404163  
VERSION AR404163.1 GI:40152203  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1503)  
AUTHORS Anderson,J.P., Basi,G., Doan,M.T., Frigon,N., John,V., Power,M.,  
Sinha,S., Tatsuno,G., Tung,J., Wang,S. and McConlogue,L.  
TITLE .beta.-secretase enzyme compositions and methods  
JOURNAL Patent: US 6627739-A 1 30-SEP-2003;  
FEATURES  
source location/Qualifiers  
1..1503  
/organism="unknown"  
/mol\_type="genomic DNA"

## ORIGIN

Alignment Scores:  
Pred. No.: 16-197  
Score: 2267.00 Length: 1503  
Percent Similarity: 99.31% Matches: 429  
Best Local Similarity: 99.08% Mismatches: 1  
Query Match: 99.08% Indels: 0  
DB: Gaps: 0

US-10-726-967a-84 (1-433) x AR404163 (1-1506)

QY 1 ThrGlnHlSGIYIleArgLeuProLeuArgSerGlyLeuGlyValAProLeuGlyLeu 20  
 Db 64 ACCGACGACGGGCACTCCGGCTGCGGAGCGGCGTGGGGGGCGCCCCCTGGGGCTG 123  
 QY 21 GluIleAsnLeuGluThrAspGluGluProGluGluProGlyValArgArgGlySerPheVal 40  
 Db 124 CGGCTGCCCCGGGAGACCGACGAGAGCCCGAGAGCCCGGCGGAGGGGCGAGCTTTGTG 183  
 QY 41 GluMetValAspAsnLeuArgGlyLysSerGlyGlnGlyTyrTyrValGluMetThrVal 60  
 Db 184 GAGATGGTGACAACTGAGGGGCAAGTCGGGGGAGGGCTCTACGCGAGATGACCGTG 243  
 QY 61 GlySerProProGlnThrLeuAsnIleLeuValAspThrGlySerSerAsnPheAlaVal 80  
 Db 244 GGCAGCCCCCGGACGAGCTCAATCATCTGGTGATACAGGACAGCACTTTCAGTG 303  
 QY 81 GluValAlaProHisProPheLeuHisArgTyrTyrGlnArgGlnLeuSerSerThrTyr 100  
 Db 304 GGTCCTCCCCCGACCCCTTCCTGCTATGCTACACAGAGGAGCGCTGTCACACATAC 363  
 QY 101 ArgAspLeuArgGlyValTyrValProTyrThrGlnGlyLysTyrProGluGlyLeu 120  
 Db 364 CGGGACCTCCGAGAGGGGTGTGTATGTGCTCTACACCGAGGGCAAGTGGAGAGGGAGCTG 423  
 QY 121 GlyThrAspLeuValSerIleProHisArgIleProAsnValThrValArgAlaAsnIleAla 140  
 Db 424 GGCACGACCTGTGAGACATCCCGCATGGCCCAAGCTCATGTGTGCTCCAACTTGCT 483  
 QY 141 AlaIleThrGluSerAspLysPhePheIleAsnGlySerAsnTrpGluGlyIleLeuGly 160  
 Db 484 GCCATCACTGAATCAGACAAAGTTCTTCATCAACGGCTCCAACTCGSAGAGCACTTCGGGG 543  
 QY 161 LeuAlaTyrAlaGluIleAlaArgProAspAspSerLeuGluProPhePheAspSerLeu 180  
 Db 544 CTGGCTTATGCTAGATTCGACAGGCTGACGACTCCCTGAGGCTTTCTTCACTCTG 603  
 QY 181 ValIleGlnThrHisValProAsnLeuPheSerLeuGlnLeuGlyValAlaGlyPhePro 200  
 Db 604 GTAAACGACACCACTTCCCAACTCTTCTCCCTGACACTTGTGTGCTGGCTTCCCC 663  
 QY 201 LeuAsnGlnSerGluValLeuAlaSerValGlyGlySerMetIleIleGlyGlyIleAsp 220  
 Db 664 CTCACACAGTCTGAAGTGTGCTGCTCTGCGAGGGAGCATATCATTTGAGAGTATCAC 723  
 QY 221 HisSerLeuTyrThrGlySerLeuTrpTyrThrProIleArgArgGluTrpTyrTyrGlu 240  
 Db 724 CACTGCTGTACACAGGCAAGTCTCTGTATACACCATCCGGCGGAGGTGATATTAG 783  
 QY 241 ValIleIleValAlaArgValGluIleAsnGlyIleAspLeuLeuMetAspCysIleAsnGluTyr 260  
 Db 784 GTGATCATTTGTGGGGTGGAGATCAATGACAGATCTGAAATATGACTCGACGAGATAC 843  
 QY 261 AsnTyrAspLysSerIleValAspSerGlyThrThrAsnLeuArgLeuProLysIleVal 280  
 Db 844 AACTATGACAAAGACATGTGTGACAGTGGCACCAACTTCGTTTCCCAAGAAAGTG 903  
 QY 281 PheGluAlaAlaValLysSerIleLeuValAlaSerSerThrGluLysPheProAspGly 300  
 Db 904 TTTGAAAGCTGCAATTCATCAAGGAGAGCTCTCCACGAGAAATCTCCGATGGT 963  
 QY 301 PheTrpLeuGlyGluGlnLeuValCysTrpGlnAlaGlyThrThrProTyrAsnIlePhe 320  
 Db 964 TTCTGGCTAGAGAGACACTGGTGTGCTGGCAGAGGCAACACCCCTTGGAACATTTTC 1023  
 QY 321 ProValIleSerLeuTyrLeuMetGlyGluValThrAsnGlnSerPheArgIleThrIle 340  
 Db 1024 CCAAGCATCTCACTTACCTATAGGGTGAAGTTCCAAACGAGTCTTCGACATCACATC 1083  
 QY 341 LeuProGlnGlnTyrLeuArgProValGluAspValAlaThrSerGlnAspAspCysTyr 360

Db 1084 CTTCGACGCAATACCTGGCGCCAGTGGAGATGTGCCACAGTCCCAAGACGATGTAC 1143  
 QY 361 LysPheAlaIleSerGlnSerSerThrGlyThrValMetGlyValAlaIleMetGlyLeu 380  
 Db 1144 AAGTTTGCCATCTCACAGATCATCCAGGGCACTGTATAGGAGCTGTATATATAGAGGGC 1203  
 QY 381 PheTyrValIlePheAspArgAlaArgLysArgIleGlyPheAlaValSerAlaCysHis 400  
 Db 1204 TTCTAGTGTCTTGTATCGGGCCGGAAGAAAGAAATGGCTTTGCTGTACAGGCTTGGCAT 1263  
 QY 401 ValHisAspGluPheArgThrAlaAlaValGluGlyProPheValThrLeuAspMetGlu 420  
 Db 1264 GTGACAGATGATTCAGAGCGAGCGGTGGAGAGCCCTTTTGTCACTTGACATGAA 1323  
 QY 421 AspCysGlyTyrAsnIleProGlnThrAspGluSerThr 433  
 Db 1324 GACTGTGCTCAACATTCCTCCACAGACAGATGATCAACC 1362

RESULT 2  
 AX700448 1506 bp DNA linear PAT 03-Apr-2003  
 LOCUS  
 DEFINITION Sequence 3 from Patent WO03012089.  
 ACCESSION AX700448  
 VERSION AX700448.1 GI:29536239  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE  
 1  
 AUTHORS  
 TITLE  
 JOURNAL  
 Patent: WO 03012089-A 3 13-FEB-2003;  
 thereof  
 FEATURES  
 location/Qualifiers  
 1..1506  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"

## ORIGIN

Alignment Scores:  
 Pred. No.: 1e-197 Length: 1506  
 Score: 2267.00 Matches: 429  
 Percent Similarity: 99.31% Conservative: 1  
 Best Local Similarity: 99.08% Mismatches: 3  
 Query Match: 99.08% Indels: 0  
 DB: 6 Gaps: 0

US-10-726-967a-84 (1-433) x AX700448 (1-1506)

QY 1 ThrGlnHlSGIYIleArgLeuProLeuArgSerGlyLeuGlyValAProLeuGlyLeu 20  
 Db 64 ACCGACGACGGGCACTCCGGCTGCGGAGCGGCGTGGGGGGCGCCCCCTGGGGCTG 123  
 QY 21 GluIleAsnLeuGluThrAspGluGluProGluGluProGlyValArgArgGlySerPheVal 40  
 Db 124 CGGCTGCCCCGGGAGACCGACGAGAGCCCGAGAGCCCGGCGGAGGGGCGAGCTTTGTG 183  
 QY 41 GluMetValAspAsnLeuArgGlyLysSerGlyGlnGlyTyrTyrValGluMetThrVal 60  
 Db 184 GAGATGGTGACAACTGAGGGGCAAGTCTGGGGGAGGGCTCTACGCGAGATGACCGTG 243  
 QY 61 GlySerProProGlnThrLeuAsnIleLeuValAspThrGlySerSerAsnPheAlaVal 80  
 Db 244 GGCAGCCCCCGGACGAGCTCAATCATCTGGTGATACAGGACAGCACTTTCAGTG 303  
 QY 81 GluValAlaProHisProPheLeuHisArgTyrTyrGlnArgGlnLeuSerSerThrTyr 100  
 Db 304 GGTCCTCCCCCGACCCCTTCCTGCTATGCTACACAGAGGAGCGCTGTCACACATAC 363



Db 304 GGTGCTGCCCCCACCCTCTTCTGATGCTACTACAGAGCAGCTGTCCAGACATAC 363  
Qy 101 ArgApPLeuAArgLygLyValTyValProTyThrgInglYlYsTPGluGlyUeu 120  
Db 364 CGGAGACTCCGGAGAGGTGTGTATGTGCTTACACCCAGGCGCAAGTGGGAAGGGGACTG 423  
Qy 121 GlyThraPLeuValSerIleProHibGlyProAenValThrValArgAlaAenIleAla 140  
Db 424 GGCACCGACTGTGTAGATGATCCCATGCGCCCAACGCTGCTGCGCTGCACATTTGCT 483  
Qy 141 AlaIleThrgIuSerAspLyPhePheIleAenGlySerAenTPGluGlyIleUeuGly 160  
Db 484 GCCATCACTGATCAGACAACTTTCTTACTACACGGCTCCAACTGGGAAGCACTCTGGGG 543  
Qy 161 LeuAlaTyAlaGluIleAlaArgProAapSerIleuGluProPheAapSerIleu 180  
Db 544 CTGGCTTAATGTGTAGATGCTGAGGCTGACGACTCCCTGGAGCTTTCTTGACTCTG 603  
Qy 181 ValIyegInThrhIsvAlProAenLeuPheSerIleuGlyUeuCyseGlyAlaGlyPhePro 200  
Db 604 GTAAAGCAGACCCACGTTCCCAACCTCTCTCCCTGAGCTTTGTGTGCTGCTGCCCTCCC 663  
Qy 201 LeuAenGInserGluValIleuAlaSerValGlyIySerMetIleIleGlyGlyIleAap 220  
Db 664 CTCAACCACTGTGAAGTGTGCTGCTGTGCGAGGAGCATGATCANTGGAGGTATCGAC 723  
Qy 221 HisSerIleuTyThrgIySerIleuTyThrProIleArgArgGluTyTyTyGlu 240  
Db 724 CACTCGCTGTACACAGGCACTCTGTGTATACACCATCGGCGGAGTGTATATATGAG 783  
Qy 241 ValIleIleValArgValGluIleAenGlyIleAapLeuIleuAapLeuIleuAapLeuIleu 260  
Db 784 GTGATCATTTGTGGGGGTGAGATCATAGCAGAGACTGAAAATGAGACTGCAGAGATAC 843  
Qy 261 AsnTyThraPlySerIleValIleAapSerGlyThrThraAenLeuArgLeuProIlyIlyVal 280  
Db 844 AACATAGACAAAGCACTGTGTGACAGTGGCACCAACCTTGTTCGCCAAGAAAGTG 903  
Qy 281 PheGluAlaIleValIySerIleValIleAaIleAaIleAaIleAaIleAaIleAaIleAa 300  
Db 904 TTGGAGCTGCACTCAATTCATCAAGGACCTCTCCACGAGAAAGTTCCCTGATGAT 963  
Qy 301 PheTPLeuGlyGluIleuValCyseTPGlnAlaGlyThrThrProTPAenIlePhe 320  
Db 964 TTCTGGCTAGAGAGCAGCTGGTGTCTGCAAGGAGGACCACTTCCTTCGATCACCATTC 1023  
Qy 321 ProValIleSerIleuTyThrMetGlyValIleThraAenGInserPheArgIleThrIle 340  
Db 1024 CCAATCATCTCACTACTAATAGGTAGGTTTACCAACCACTCTTCGATCACCATTC 1083  
Qy 341 LeuProGInglTyThraArgProValGluAapValAlaThraSerGlyAapAapCyseTy 360  
Db 1084 CTTCCGACAAATACCTGGCGCAGTGAAGATGTGGCCACTCCCAAGACGACTGTTC 1143  
Qy 361 IySPheAlaIleSerGInserSerThrglyThrValMetGlyAlaValIleMetGluGly 380  
Db 1144 AAGTTTGCATCTCAAGTCACTCAAGGCGCACTGTATAGGAGCTGTATCATGAGAGGC 1203  
Qy 381 PheTyValValPheAapArgAlaArgIySAArgIleGlyPheAlaValSerAlaCyseHis 400  
Db 1204 TTCTAAGTTTGTCTTGTATCGGGCCGAAAGCAATGTGCTTGTCTGTCAGGCTTGCAT 1263  
Qy 401 ValHisAapGluPheArgThraAlaAlaValGluGlyPropheValThraLeuAapMetGlu 420  
Db 1264 GTGCACGATGAGTTCAGAGCGGACGGTGGAGGCTTTTGTGCACCTTGGACATGGA 1323  
Qy 421 AapCyseGlyTyThraIleProGInThraPgluSerThr 433  
Db 1324 GACTGTGCTACAACTTCCACAGCAGATGATGATCAAC 1362  
RESULT 4  
BD103723.

LOCUS BD103723 1527 bp DNA linear PAT 27-AUG-2002  
DEFINITION Beta secretase inhibitor.  
ACCESSION BD103723  
VERSION BD103723.1 GI:22649297  
KEYWORDS WO 0187293-A/6.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1527)  
AUTHORS Miyamoto,M., Matsui,J., Fukumoto,H. and Tarui,N.  
TITLE Beta secretase inhibitor  
JOURNAL Patent: WO 0187293-A 6 22-NOV-2001;  
TAKEDA CHEMICAL INDUSTRIES LTD,MASAO MIYAMOTO,JUNJI MATSUI,  
HIROAKI FUKUMOTO,NAOKI TARUI  
COMMENT OS Homo sapiens (human)  
PN WO 0187293-A/6  
PD 22-NOV-2001  
PF 18-MAY-2001 WO 2001JP004144  
PR 19-MAY-2000 JP 00P 152758  
PI MASAO MIYAMOTO,JUNJI MATSUI, HIROAKI FUKUMOTO,NAOKI TARUI  
PC 4164' A61K31/00,A61P25/00,A61P25/16,A61P25/28  
CC Beta secretase inhibitor  
FH Key location/Qualifiers  
FT source 1..1527  
location/Qualifiers  
1..1527  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.02e-197 Length: 1527  
Score: 2267.00 Matches: 429  
Percent Similarity: 99.31% Conservative: 1  
Best Local Similarity: 99.08% Mismatches: 3  
Query Match: 99.08% Indels: 0  
Gaps: 0  
US-10-726-967a-84 (1-433) x BD103723 (1-1527)  
Qy 1 ThrGlnHisGlyIleArgLeuProLeuArgSerGlyLeuGlyValAProLeuGlyLeu 20  
Db 64 ACCACACAGGCACTCCGCTGCCCTGCGCAGCGGCTGGGGGGGCCCCCTGGGGCTG 123  
Qy 21 GluIleAenLeuGluThraPgluGluProGluGluProGlyAArgGlySerPheVal 40  
Db 124 CGGCTGCCCCGGAGAGCCAGCAGAAAGCCCGAGAGCCCGCGGAGGGGAGCTTTGTG 183  
Qy 41 GluMetValAspAsnLeuArgGlyIySerGlyGlnGlyTyTyTyValGluMetThraVal 60  
Db 184 GAGATGTGTGAACACTGAGGGGCAAGTCGGGGCAGGGCTTACTAGTGAATGACCTG 243  
Qy 61 GlySerProProGlnThraAenIleLeuValAapThrgIySerSerAsnPheAlaVal 80  
Db 244 GGCAGCCCCCAGCAGCTCAACATCTGTTGATACAGCAGCAGCACTTGTGACGTG 303  
Qy 81 GlyAlaAlaProHibProPheLeuHibArgTyTyTyGlnArgGlnLeuSerSerThrTy 100  
Db 304 GGTGCTGCCCCCACCCTCTCTGATGCTACTACAGAGCAGCTGTCCAGACATAC 363  
Qy 101 ArgApPLeuAArgLygLyValTyValProTyThrgInglYlYsTPGluGlyUeu 120  
Db 364 CGGAGACTCCGGAGAGGTGTGTATGTGCTTACACCCAGGCGCAAGTGGGAAGGGGACTG 423  
Qy 121 GlyThraPLeuValSerIleProHibGlyProAenValThrValArgAlaAenIleAla 140  
Db 424 GGCACCGACTGTGTAGATGATCCCATGCGCCCAACGCTGCTGCGCTGCACATTTGCT 483



QY 141 AlaIlethrgluserAspLysePhePheIleanglyseAsnTrpGluGlyIleleugly 160  
DB 484 GCCATCACTGATCAGACAAAGTTCTTCAATCAACGGCTCCAACTGGAGAGGCACTCGGG 543  
QY 161 LeuAlaYrAlaGluIleAlaPProaPAspSerLeuGluPProPhePheAspSerLeu 180  
DB 544 CTGGCCTATGCTGAATGCTCCAGGCTTCAAGCACTCCCTGGAGCTTTCTTGAATCTCTG 603  
QY 181 ValLyseGlnThrHisValProaPLeuPheSerLeuGlnLeuCyseGlyValaGlyPhePro 200  
DB 604 GTPAAGCAGACCAAGTCCCAACCTCTTCTCCCTGACAGCTTGGTGGCTGGCTTCC 663  
QY 201 LeuAsnGlnSerGluValLeuAlaSerValaGlyGlySerMetIleIleGlyIleLeuP 220  
DB 664 CTCACCAAGTCAAGTGTGGCTGCTGTGCGAGGAGATGATATCAATGGAGATTCGAC 723  
QY 221 HisSerLeuYrThrGlySerLeuTrpYrThrProIleAspArgGluTrpYrYrGlu 240  
DB 724 CACTGCTGTACACAGCAGCTCTGCTGTATACCCATCCGCGGGAGTGGTATTTAG 783  
QY 241 ValIleIleValaArgValaGluIleAnglyGlnAspLeuLyseAspCyseLyseGlyYr 260  
DB 784 GTGATCATTTGTGGGGTGGAGATCAATGACAGGATCTGAATGACCTGCAAGAGGTAC 843  
QY 261 AsnYrAspLyseSerIleValaPserGlyYrThrAsnLeuArgLeuProLyseVala 280  
DB 844 AACTATGACAAAGCAATGTGGCAGTGGCACCAACCACTTCGTTGGCCAAAGAAAG 903  
QY 281 PheGluAlaAlaValLyseSerIleValaAlaSerSerThrGluLysePheProAspGly 300  
DB 904 TTTGAAGCTGACGATCAATTCATCAAGGACGCTCTCCACGAGAAAGTTCCTATGAT 963  
QY 301 PheTrpLeuGlyGluGlnLeuValCyseTrpGlnAlaGlyYrThrProTrpAsnIlePhe 320  
DB 964 TTCTGCTAGAGAGACAGCTGTGTGCTGTGGCAAGGACCAACCTTGGAAACATTTTC 1023  
QY 321 ProValIleSerLeuYrLeuMetGlyGluValaThrAsnGlnSerPheArgIleThrIle 340  
DB 1024 CCAATCATCTCACTCACTCAATGGGTGAGTTACCAACCAAGCTTCCTGGCATCACATC 1083  
QY 341 LeuProGlnGlyTrpLeuArgProValaGluAspValaAlaThrSerGlnAspAspCyseYr 360  
DB 1084 CTTCCGACCAAACTGCGGCGAGTGAAGATGTGGCCAGCTCCAAAGCACTTTAC 1143  
QY 361 LysPheAlaIleSerGlnSerSerThrGlyYrValaMetGlyAlaValIleMetGluGly 380  
DB 1144 AAGTTGCCATCTCAAGTCATCCAGGCACTGTTATGGAGCTTTATCATGAGAGGG 1203  
QY 381 PheYrValaValaPheAspArgAlaArgLyseArgIleGlyPheAlaValaSerAlaCyseHis 400  
DB 1204 TTCTACGTTGTTCTTATCGGGCCGAAAAAGAAATTGGCTTGTCTGACGCGCTTCCAT 1263  
QY 401 ValHisAspGluPheArgThrAlaAlaValaGluGlyProPheValaThrLeuAspMetGlu 420  
DB 1264 GTGCAAGATGATTCAGAGAGGAGCGGGTGAAGGCCCTTTTGTCACTTGGACATGAA 1323  
QY 421 AspCyseGlyYrAsnIleProGlnThrAspGluSerThr 433  
DB 1324 GACTGTGGCTAACAATTCCACAGACAGATGATCAACC 1362  
RESULT 5  
AR224093 2070 bp DNA linear PAT 26-SBP-2002  
LOCUS DEFINITION  
AR224093 Sequence 3 from patent US 6440698.  
ACCESSION AR224093  
VERSION AR224093.1 GI:23332753  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 2070)  
AUTHORS Gurney,M.E., Bienkowski,M.J., Heinrikson,R.L., Parodi,L.A. and Yan,R.

TITLE Alzheimer's disease secretase, APP substrates therefor, and uses therefor  
JOURNAL Patent: US 6440698-A 3 27-AUG-2002;  
FEATURES Location/Qualifiers  
SOURCE 1..2070  
/organism="unknown"  
/mol\_type="genomic DNA"  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.5e-197 Length: 2070  
Score: 2267.00 Matches: 429  
Percent Similarity: 99.31% Conservative: 1  
Best Local Similarity: 99.08% Mismatches: 3  
Query Match: 99.08% Indels: 0  
DB: 6 Gaps: 0  
US-10-726-967A-84 (1-433) x AR224093 (1-2070)  
QY 1 ThrGlnHisGlyIleArgLeuProLeuArgSerGlyLeuGlyValaProLeuGlyLeu 20  
DB 64 ACCCAGCAGGCAATCGGCTGCTGCGACGGCTTGGGGGCGCCCCCTGGGGCTG 123  
QY 21 GluIleAsnLeuGluThrAspGluGluProGluGluProGlyArgArgGlySerPheVal 40  
DB 124 CGGCTGCCCGGGAGACCGAGAAAGCCGAGAGCCCGGCGGAGGGGAGCTTGTG 183  
QY 41 GluMetValaAspAsnLeuArgGlyLyseSerGlyGlnGlyYrYrYrValaGluMetThrVal 60  
DB 184 GAGATGTGTGACAACTGAGGGGCAAGTGGGGGCGGCTACTACGTGAGATGACCGTG 243  
QY 61 GlySerProProGlnThrLeuAsnIleLeuValaAspThrGlySerSerAspPheAlaVal 80  
DB 244 GGCAGCCCCCGCAACCTCAATCCGTGGATGACGAGCAAGCAATTAATTTGCACTG 303  
QY 81 GlyAlaAlaProHisProPheLeuHisArgYrYrYrGlnArgGlnLeuSerSerThrYr 100  
DB 304 GGTGCTGCCCGCCACCCCTCTCTGATGCTACTACAGAGCACTGTCCAGACATAC 363  
QY 101 ArgAspLeuArgLyseGlyValaYrValaProLyThrGlnGlyLyseTrpGluGlyLeu 120  
DB 364 CGGACCTCCGGAAGGTGTGTATGTGCTTACACCCAGGGCAAGTGGGAAGGGAGCTG 423  
QY 121 GlyThrAspLeuValSerIleProHisGlyProaPValaThrValArgAlaAsnIleAla 140  
DB 424 GGCACCGACTGTGTAAGCATCCCAATGCCCCCAAGTCATCTGTGCTCCAACTTGT 483  
QY 141 AlaIlethrgluserAspLysePhePheIleanglyseAsnTrpGluGlyIleleugly 160  
DB 484 GCCATCACTGATCAGACAAAGTTCTTCAATCAACGGCTCCAACTGGAGAGGCACTCGGG 543  
QY 161 LeuAlaYrAlaGluIleAlaPProaPAspSerLeuGluPProPhePheAspSerLeu 180  
DB 544 CTGGCCTATGCTGAATGCTCCAGGCTTCAAGCACTCCCTGGAGCTTTCTTGAATCTCTG 603  
QY 181 ValLyseGlnThrHisValProaPLeuPheSerLeuGlnLeuCyseGlyValaGlyPhePro 200  
DB 604 GTPAAGCAGACCAAGTCCCAACCTCTTCTCCCTGACAGCTTGGTGGCTGGCTTCC 663  
QY 201 LeuAsnGlnSerGluValLeuAlaSerValaGlyGlySerMetIleIleGlyIleLeuP 220  
DB 664 CTCACCAAGTCAAGTGTGGCTGCTGTGCGAGGAGATGATATCAATGGAGATTCGAC 723  
QY 221 HisSerLeuYrThrGlySerLeuTrpYrThrProIleAspArgGluTrpYrYrGlu 240  
DB 724 CACTGCTGTACACAGCAGCTCTGCTGTATACCCATCCGCGGGAGTGGTATTTAG 783  
QY 241 ValIleIleValaArgValaGluIleAnglyGlnAspLeuLyseAspCyseLyseGlyYr 260  
DB 784 GTGATCATTTGTGGGGTGGAGATCAATGACAGGATCTGAATGACCTGCAAGAGGTAC 843  
QY 261 AsnYrAspLyseSerIleValaPserGlyYrThrAsnLeuArgLeuProLyseVala 280

Db 844 AACTATGACAAAGCATTTGTGGACAGTGGACACCAACCTTGCTTGCCCAAGAAAGTG 903  
Qy 281 PheGluAlaAlaValIleYserIleYsaIaIaIseSerThrGluYsPheProAspGly 300  
Db 904 TTGTAAGCTGCAGTCAAATTCATCAAGGACAGCTCTCCACGAGAAAGTTCCCGATGGT 963  
Qy 301 PheTrpLeuGlyGluGlnLeuValCysTrpGlnIleValIleGlyThrThrProTrpAsnIlePhe 320  
Db 964 TTCTGGCTAGAGAGACAGCTGTGTGTGTGGCAAGCAGACACCCCTTGGAACATTTTC 1023  
Qy 321 ProValIleSerLeuYrLeuMetGlyGluValThrAsnGlnSerPheArgIleThrIle 340  
Db 1024 CCAGCATCTCACTCACTCACTAATGAGGTTCACCAACAGCTCTCCGATCAGCATC 1083  
Qy 341 LeuProGlnGlnIleYrLeuArgProValGluAspValIaThrSerGlnAspAspCysTrp 360  
Db 1084 CTTCCGACGAAATCTCGCGGCGAGTGAAGATGTGCCACAGTCCCAAGACGACTGTAC 1143  
Qy 361 LysPheAlaIleSerGlnSerSerThrGlyThrValMetGlyAlaValIleMetGluGly 380  
Db 1144 AAGTTTGCATCTCAAGCTATCCACGGGCACTGTTATGGAGCTGTATCATGAGGGC 1203  
Qy 381 PheTrpValValPheAspArgAlaArgLysArgIleGlyPheAlaValSerAlaCysHis 400  
Db 1204 TTCTACGTTGCTTTGATCGGGCCCGAAACGAATTGGCTTGCTGTCAAGCCGCTTGCAT 1263  
Qy 401 ValHisAspGluPheArgThrAlaAlaValGluGlyProPheValThrLeuAspMetGlu 420  
Db 1264 GTGCACATAGATTCAGAGACGGACGGGTGAAAGGCCCTTTGTCACTTGACATGAGAA 1323  
Qy 421 AspCysGlyYrAsnIleProGlnThrAspGluSerThr 433  
Db 1324 GACTGTGGCTACAAACATTCACAGACAGATGATCAACC 1362

RESULT 6  
LOCUS AR269224 2070 bp DNA linear PAT 10-APR-2003  
DEFINITION Sequence 3 from patent US 650667.  
ACCESSION AR269224  
VERSION AR269224.1 GI:29700192  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE Unpublished.  
1 (bases 1 to 2070)  
AUTHORS Gurney,M.E., Bienkowi,M.J., Heinrichson,R.L., Parodi,L.A. and Yan,R.  
TITLE Aspartyl protease 2 (Asp2) antisense oligonucleotides  
JOURNAL Patent: US 650667-A 3 31-DEC-2002;  
FEATURES Location/Qualifiers  
source 1..2070  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Alignment Scores:  
Pred. No.: 1,5e-197 Length: 2070  
Score: 2267.00 Matches: 429  
Percent Similarity: 99.31% Conservative: 1  
Best Local Similarity: 99.08% Mismatches: 3  
Query Match: 99.08% Indels: 0  
DB: 6 Gaps: 0  
US-10-726-967a-84 (1-433) x AR269224 (1-2070)

Qy 1 ThrGlnHISGlyIleArgLeuProLeuArgSerGlyLeuGlyValAlaProLeuGlyLeu 20  
Db 64 ACCGACGACGGGATCCGGCTGCCCTGCGACGGCCCTGGGGGGCGCCCCCTGGGGCTG 123  
Qy 21 GluIleAsnLeuGluThrAspGluGluProGluGluProGlyArgArgGlySerPheVal 40  
Db 124 CGGCTGCCCGGGAGACCGACGAAGAGCCCGAGAGAGCCGGCCGGAAGGGCAGCTTGTG 183

Qy 41 GluMetValAspAsnLeuArgGlyLysSerGlyGlnGlyYrYrYrValGluMetThrVal 60  
Db 184 GAGATGTGTGACAACTGAGGGGCAAGTCGGGGCAAGGCTCTACGTGGAATGACCTGTG 243  
Qy 61 GlySerProProGlnThrLeuAsnIleLeuValAspTrpGlyYserSerAsnPheAlaVal 80  
Db 244 GGCAGCCCCCGCAGACCTCAACATCTGTGTGATACAGGACACAGTAACCTTTGCGAGTG 303  
Qy 81 GluAlaAlaProHisAspPheLeuHisAspGlyYrYrYrGlnArgGlnLeuSerSerThrYr 100  
Db 304 GGTGTGCCCCCACCCTTCTGTGATGTCTACTACCAAGAGGAGCTGTCCAGCACATAC 363  
Qy 101 ArgAspLeuArgLysGlyValYrValProTrpThrGlnGlyLysTrpGluGluLeu 120  
Db 364 CGGAGCTCCGGAGAGGGGTGTATGTGCTTACACCCAGGCGAAGTGGAAAGGGAGCTG 423  
Qy 121 GlyThrAspLeuValSerIleProHisGlyProAsnValThrValArgAlaAsnIleAla 140  
Db 424 GGCACCGACCTGGTAAGCATCCCCCATGGCCCCCAACGCTCACTGTGCGCAACATTTGCT 483  
Qy 141 AlaIleThrGlnSerAspLysPhePheIleAsnGlySerAsnTrpGlnGlyIleLeuGly 160  
Db 484 GGCATCACTGAATACAGACAACTTTCTATCAACGCTTCAACTGGGAAGGCATCTGGGG 543  
Qy 161 LeuAlaTrpAlaGluIleAlaArgProAspAspSerLeuGluProPheAspSerLeu 180  
Db 544 CTGGCTATGTGTGAATGTGGCAGGCTGACGACATCTCTGAGCTTCTTGACTCTCTG 603  
Qy 181 ValLysGlnThrHisValProAsnLeuPheSerLeuGlnLeuGlyValAlaGlyPhePro 200  
Db 604 GTAAAGACAGACCCAGCTTCCCAACCTTTCTCCCTGAGCTTGTGTGTGCTTCCCC 663  
Qy 201 LeuAsnGlnSerGluValLeuAlaSerValGlyYserMetIleIleGlyIleAsp 220  
Db 664 CTCAACAGCTGAAGGTGCTGGCTGTGTGAGGAGATGATTCATTGAGATACGAC 723  
Qy 221 HisSerLeuYrThrGlySerLeuTrpYrThrProIleArgArgGluTrpYrYrGlu 240  
Db 724 CACTGGCTGTACAGGCACTCTGTGTATACACCATCCGGGGGAGGTGATATATGAG 783  
Qy 241 ValIleIleValArgValGluIleAsnGlyGlnAspLeuYsMetAspCysLysGluYr 260  
Db 784 GTCATCATTTGTGGGGTGGAGATCATGACAGATCTBAATGTGACTGCAAGAGATC 843  
Qy 261 AsnTrpAspLysSerIleValAspSerGlyThrThrAsnLeuArgLeuProLysVal 280  
Db 844 AACTATGACAAAGCATTTGTGGACAGTGGCACACCAACCTTGTTGCCCAAGAAAGTG 903  
Qy 281 PheGluAlaAlaValIleYserIleYsaIaIaIseSerThrGluYsPheProAspGly 300  
Db 904 TTGTAAGCTGCAGTCAAATTCATCAAGGACAGCTCTCCACGAGAAAGTTCCCGATGGT 963  
Qy 301 PheTrpLeuGlyGluGlnLeuValCysTrpGlnIleValIleGlyThrThrProTrpAsnIlePhe 320  
Db 964 TTCTGGCTAGAGAGACAGCTGTGTGTGTGGCAAGCAGACACCCCTTGGAACATTTTC 1023  
Qy 321 ProValIleSerLeuYrLeuMetGlyGluValThrAsnGlnSerPheArgIleThrIle 340  
Db 1024 CCAGCATCTCACTCACTCACTAATGAGGTTCACCAACAGCTCTCCGATCAGCATC 1083  
Qy 341 LeuProGlnGlnIleYrLeuArgProValGluAspValIaThrSerGlnAspAspCysTrp 360  
Db 1084 CTTCCGACGAAATCTCGCGGCGAGTGAAGATGTGCCACAGTCCCAAGACGACTGTAC 1143  
Qy 361 LysPheAlaIleSerGlnSerSerThrGlyThrValMetGlyAlaValIleMetGluGly 380  
Db 1144 AAGTTTGCATCTCAAGCTATCCACGGGCACTGTTATGGAGCTGTATCATGAGGGC 1203  
Qy 381 PheTrpValValPheAspArgAlaArgLysArgIleGlyPheAlaValSerAlaCysHis 400  
Db 1204 TTCTACGTTGCTTTGATCGGGCCCGAAACGAATTGGCTTGCTGTCAAGCCGCTTGCAT 1263  
Qy 401 ValHisAspGluPheArgThrAlaAlaValGluGlyProPheValThrLeuAspMetGlu 420

Db 1264 GTGACGAGTGAAGTTCAGGACGGCAGCGGGAAGGCCCTTTGTGACCTTGACATGGAA 1323  
Qy 421 AAPPYSGLYTYRASNILEPROGLINTHRASPGLUSERTHR 433  
Db 1324 GACTGTGGCTACACATTCCTCCACAGACGATGAGTCAACC 1362

RESULT 7  
AR478778 2070 bp DNA linear PAT 14-MAY-2004  
LOCUS AR478778  
DEFINITION Sequence 3 from patent US 6699671.  
ACCESSION AR478778  
VERSION AR478778.1 GI:47237498  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 2070)  
AUTHORS Gurney,M.E., Bienkowski,M.J., Heinrichson,R.L., Parodi,L.A. and  
Van,R.  
TITLE Alzheimer's disease secretase, APP substrates therefor, and uses  
therefor  
JOURNAL Patent: US 6699671-A 3 02-MAR-2004;  
FEATURES Location/Qualifiers  
source 1..2070  
/organism="unknown"  
/mol\_type="genomic DNA"

ALIGNMENT Scores:  
Pred. No.: 1.5e-197 Length: 2070  
Score: 2267.00 Matches: 429  
Percent Similarity: 99.31% Conservative: 1  
Best Local Similarity: 99.08% Mismatches: 3  
Query Match: 99.08% Indels: 0  
DB: 6 Gaps: 0

US-10-726-967a-84 (1-433) x AR478778 (1-2070)

Qy 1 ThrGlnHlgLy1LeArgLeuProLeuArgSerGlyLeuGlyValAlaProLeuGlyLeu 20  
Db 64 ACCACACAGGACATCCGGCTGCGGCTGCGAGCGCTGGGGGGCCCCCTGGGGCTG 123  
Qy 21 GluileAsnLeuGlyLthrAspGluGluProGluGluProGlyValArgArgGlySerPheVal 40  
Db 124 CGGCTGCCCGGAGACCGACGAGAGCCGAGAGACCCGCGGCGGAGGCGGACCTTGTG 183  
Qy 41 GluMetValAspAsnLeuArgGlyLysSerGlyGlnGlyTyrTyrValGluMetThrVal 60  
Db 184 GAGATGCTGACCACTGAGGGGCAAGTCGGGCGAGGCTACTACGTGAGATGACCGTG 243  
Qy 61 GlySerProProGluInthrLeuAsnIleLeuValAspThrGlySerSerAsnPheAlaVal 80  
Db 244 GGCAGCGCCCGGAGAGCTCAACATCTGGTGATACAGGCGACGTAACCTTGCAGTG 303  
Qy 81 GlyAlaAlaProHisProPheLeuHISArgTyrTyrGlnArgGlnLeuSerSerThrTyr 100  
Db 304 GGTGTGCCCCCACCCTTCTCGCATCGCTACTACAGAGGAGCGTGTCCAGACATAC 363  
Qy 101 ArgAspLeuArgGlyValTyrValProTyrThrGlnGlyLysTrpGluGlyLeu 120  
Db 364 CGGAGCTCCGGAGAGGTGTGTGTGCTTACACCCAGGCGCAAGTGGGAGGGAGCTG 423  
Qy 121 GlyThrAspLeuValSerIleProHisGlyProAsnValThrValArgAlaAsnIleAla 140  
Db 424 GGCACGACCTGGTGAAGCATCCCGCATGCCCCCAAGCTCACTGTGCTGCCAATTGCT 483  
Qy 141 AlaIleThrGluSerAspLysPhePheIleAsnGlySerAsnTrpGluGlyIleLeuGly 160  
Db 484 GCCATCACTGAATCAGACAAAGTTCTTTCATCAAGGCTCCAACTGGGAGGACATCGGG 543  
Qy 161 LeuAlaTyrAlaGluIleAlaArgProAspAspSerLeuGluProPheIleAspSerLeu 180

Db 544 CTGGCTATGCTGAGATTCAGGCTGACGATCCCTCGAGCCCTTTCTTGTACTCTG 603  
Qy 181 ValLysGlnInthrHisValProAsnLeuPheSerLeuGlnLeuGlyValAlaGlyPhePro 200  
Db 604 GTAAAGGAGACCCAGCTTCCCACTCTTCCCTGACGCTTGTGGCTTGGCTTCCC 663  
Qy 201 LeuAsnGlnSerGluValLeuAlaSerValGlyLysSerMetIleIleGlyIleAsp 220  
Db 664 CTCACCGCTGTGAAGTCTGCTGCTCGGAGGACATGATCATTTGAGATGAC 723  
Qy 221 HisSerLeuTyrThrGlySerLeuTrpTyrThrProIleArgArgGluTrpTyrGlu 240  
Db 724 CACTGCTGTACAGCGGAGCTCTGATATACCCATCCGCGGAGTGTGATATATAG 783  
Qy 241 ValIleIleValArgValGluIleAsnGlyGlnAspLeuLysMetAspCysLysGlyLys 260  
Db 784 GTCATCATTTGGCGGTGAGATCAATGACAGGATTCGAAATGAGTGCAGCGAGGATAC 843  
Qy 261 AsnTyrAspLysSerIleValAspSerGlyThrThrAsnLeuArgLeuProLysIleVal 280  
Db 844 AACTATGACAGAGCATTTGTGACAGTGGCACACCAACCTTCTGCTGCCAAGAACTG 903  
Qy 281 PheGluAlaAlaValLysSerIleLysAlaAspSerThrGluLysPheProAspGly 300  
Db 904 TTTGAAGCTGACGATCAATTCATCAAGGACGCTCTCCAGAGAAAGTTCCTGATGCT 963  
Qy 301 PheTrpLeuGlyGlyGlnLeuValCysTrpTrpGlnAlaGlyThrThrProTrpAsnIlePhe 320  
Db 964 TTCTGCTTGAAGAGACAGCTGGTGTGCTGGCAGAGGACACACCCCTTGAAACATTTTC 1023  
Qy 321 ProValIleSerLeuTyrLeuMetGlyGlyValThrAsnGlnSerPheArgIleThrIle 340  
Db 1024 CCAGTCATCTCACTTACTTAATGGGTAGAGTTACCAACCACTCTTCCGATCACATC 1083  
Qy 341 LeuProGlnGlnTyrLeuArgProValGluAspValAlaThrSerGlnAspAspCysTyr 360  
Db 1084 CTTCGCGAGCAATACCTCGGCGCAGTGGAAGTGTGGCCAGTCCCAAGACACATGTATC 1143  
Qy 361 LysPheAlaIleSerGlnSerSerThrGlyThrValMetGlyAlaValIleMetGlyGly 380  
Db 1144 AAGTTGGCATCTCAAGATTCACAGGCGACCTGTATAGGAGCTGTATCATGAGAGGC 1203  
Qy 381 PheTyrValAlaPheAspArgAlaArgLysArgIleGlyPheAlaValSerAlaCysHis 400  
Db 1204 TTCTACGTGTCTTGTATCGGCCCGCAAAACAAATGGCTTTGCTGACGCTTGCCAT 1263  
Qy 401 ValHisAspGluPheArgThrAlaAlaValGluGlyProPheValThrLeuAspMetGlu 420  
Db 1264 GTGACGAGTGAATTCAGAGCGGCAAGCGGTGGAAGGCCCTTTGTCACTTGACATGMA 1323  
Qy 421 AAPPYSGLYTYRASNILEPROGLINTHRASPGLUSERTHR 433  
Db 1324 GACTGTGGCTACACATTCACAGACGATGAGTCAACC 1362

RESULT 8  
AR487345 2070 bp DNA linear PAT 14-MAY-2004  
LOCUS AR487345  
DEFINITION Sequence 3 from patent US 6706485.  
ACCESSION AR487345  
VERSION AR487345.1 GI:47252443  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 2070)  
AUTHORS Gurney,M.E., Bienkowski,M.J., Heinrichson,R.L., Parodi,L.A. and  
Van,R.  
TITLE Method of identifying agents that inhibit APP processing activity  
JOURNAL Patent: US 6706485-A 3 16-MAR-2004;  
FEATURES Location/Qualifiers  
source 1..2070  
/organism="unknown"  
/mol\_type="genomic DNA"

## ORIGIN

## Alignment Scores:

Pred. No.:	1.5e-197	Length:	2070
Score:	2267.00	Matches:	429
Percent Similarity:	99.31%	Conservative:	1
Best Local Similarity:	99.08%	Mismatches:	3
Query Match:	99.08%	Indels:	0
DB:	6	Gaps:	0

US-10-726-967a-84 (1-433) x AR487345 (1-2070)

```

QY 1 ThrGlnHisGlyIleArgLeuProLeuArgSerGlyLeuGlyValProLeuGlyLeu 20
Db ACCGACGACGGGCACTCCGGCTCCCTGGCGACGGGCTGGGGGGGCGCCCTGGGGCTG 123
QY 21 GluIleAsnLeuGluThrAspGluGluProGluGluProGlyArgArgGlySerPheVal 40
Db CGGCTGCCCGGGAGACCGACGAAGAGCCCGAGGAGCCGGCGGAGGGGCACTTTGTG 183
QY 41 GluMetValAspAsnLeuArgGlyLysSerGlyGlnGlyTyrTyrValGluMetThrVal 60
Db GAGATGTGACCACTCGAGGGGCAAGTCGAGGCGAGGCTTACCTGAGATGACCGCTG 243
QY 61 GlySerProProGlnThrLeuAsnIleLeuValAspThrGlySerSerAsnPheAlaVal 80
Db GGCAGCCCCCGGAGACGCTCAACATCTGTGGATACAGCAGCATTACTTGGAGTG 303
QY 81 GlyAlaIleAlaProHisProPheLeuHisArgTyrTyrGlnArgGlnLeuSerSerThrTyr 100
Db GGGTGTGCCCCCGACCCCTTCTGCACTGCTACTACAGGAGCGCTGTCACAGCATATAC 363
QY 101 ArgAspLeuArgLysGlyValTyrValProTyrThrGlnGlyLysThrGluGlyGluLeu 120
Db CGGGAGCTCCGGAGAGGTGTGTATGTGCTTACACCCAGGGCAAGTGGAGAGGGAGCTG 423
QY 121 GlyThrAspLeuValSerIleProHisGlyProAsnValThrValArgAlaAsnIleAla 140
Db GGCACGCACTCGGTAGCATCCCCCAGGGCCCAAGCTGCTGTGCTGCCAACAATTGCT 483
QY 141 AlaIleThrGlySerAspLysPhePheIleAsnGlySerAsnThrGluGlyIleLeuGly 160
Db GCCATCACTGATCAACAGAGTTCTTCATCAACGCGCTCCAACTGGAGAGGCACTCTGGGG 543
QY 161 LeuAlaTyrAlaGluIleAlaArgProAspAspSerLeuGluProPhePheAspSerLeu 180
Db CTGGCTATGTGTAGATGCGAGGCTGACGACTCCCTGGAGCTTCTTTTACTCTCTG 603
QY 181 ValIleGlnThrHisValProAsnLeuPheSerLeuGlnLeuGlyAlaGlyPhePro 200
Db GTAAAGCAGACCCAGCTGCCCAACCTTCTCCCTGCACTTGTGTGCTGGCTTCCCC 663
QY 201 LeuAsnGlnSerGlyValLeuAlaSerValGlyGlySerMetIleGlyGlyIleAsp 220
Db CTCACACCACTGAGAGTGTGCTGCTGTGCGAGGGAGCATATCATTTGGAGGATTCAGC 723
QY 221 HisSerLeuTyrThrGlySerLeuTyrTyrThrProIleArgArgGluTyrTyrGlu 240
Db CACTGCGCTACACAGCGCAGTCTGTGTATACACCACTCCGGCGGAGTGTGATTATAG 783
QY 241 ValIleIleValArgValGluIleAsnGlyGlnAspLeuLysMetAspCysLysGluTyr 260
Db GTCACTATTGTGCGGGTGGAGATCATGACAGATCGAATAATGACTCGCAAGAGATAC 843
QY 261 AsnTyrAspLysSerIleValAspSerGlyThrThrAsnLeuArgLeuProLysIleVal 280
Db AACCTATGACAAAGACATTGTGGACAGTGGACACCAACCTTCTGTTGCCCAAGAAAGTG 903
QY 281 PheGluIleAlaValLysSerIleLysAlaAlaSerSerThrGluLysPheProAspGly 300
Db TTTGAAGCTGCAAGTCAAAATTCATCAAGGACAGCTCTTCCACGAGAAAGTTCCCTGATG 963
QY 301 PheThrLeuGlyGluGlnLeuValCysThrGlnAlaGlyThrThrProThrAsnIlePhe 320

```

```

Db 964 TTCTGCTAGAGACAGCTGGTGTGTGGACAGGACACACCTTGGAACTTTTC 1023
QY 321 ProValIleSerLeuTyrLeuMetGlyValThrAsnGlnSerPheArgIleThrIle 340
Db 1024 CCAAGTCACTTCACTTACTTAATGGGTATGATCAACAGCATGCTTCCGATCACCATC 1083
QY 341 LeuProGlnGlnTyrLeuArgProValGluAspValAlaThrSerGlnAspAspCysTyr 360
Db 1084 CTTCCGACAGCAATACCTCGGCGAGTGGAAAGATGTGGCAGCTCCCAAGAGCATGTTAC 1143
QY 361 LysPheAlaIleSerGlnSerSerThrGlyThrValMetGlyAlaValIleMetGluGly 380
Db 1144 AAGTTTGCATCTCACTCATTCACAGGCGACTGTATAGGAGCTGTATCATGAGAGGC 1203
QY 381 PheTyrValAlaPheAspArgAlaArgLysArgIleGlyPheAlaValSerAlaCysHis 400
Db 1204 TTCTACGTTGTCTTTGATCGGGCCCGAAGAACAAATGGCTTGTGTCAAGGCTTGGCAT 1263
QY 401 ValHisAspGluPheArgThrAlaAlaValGluGlyProPheValThrLeuAspMetGlu 420
Db 1264 GTGCACGATGAGTTCAGGACGGCAGCGGTGGAGGCCCTTTTGTCACTTGGACATGGA 1323
QY 421 AspCysGlyTyrAsnIleProGlnThrAspGluSerThr 433
Db 1324 GACTGTGGCTCAACATTCACAGACAGATGATCAACC 1362

```

## RESULT 9

AR531984

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

ORIGIN

Alignment Scores:

Pred. No.:	1.5e-197	Length:	2070
Score:	2267.00	Matches:	429
Percent Similarity:	99.31%	Conservative:	1
Best Local Similarity:	99.08%	Mismatches:	3
Query Match:	99.08%	Indels:	0
DB:	6	Gaps:	0

US-10-726-967a-84 (1-433) x AR531984 (1-2070)

```

QY 1 ThrGlnHisGlyIleArgLeuProLeuArgSerGlyLeuGlyValProLeuGlyLeu 20
Db 64 ACCGACGACGGGATCCGGCTCCCTGGCGACGGGCTGGGGGGGCGCCCTGGGGCTG 123
QY 21 GluIleAsnLeuGluThrAspGluGluProGluGluProGlyArgArgGlySerPheVal 40
Db 124 CGGCTGCCCGGGAGACCGACGAAGAGCCCGAGGAGCCCGGAGGGGAGGAGCTTTGTG 183
QY 41 GluMetValAspAsnLeuArgGlyLysSerGlyGlnGlyTyrTyrValGluMetThrVal 60
Db 184 GAGATGTGACCACTCGAGGGGCAAGTCGAGGCGGCTTACTAGTGAATGACCGTG 243
QY 61 GlySerProProGlnThrLeuAsnIleLeuValAspThrGlySerSerAsnPheAlaVal 80

```

Db 244 GGACAGCCCCCGACAGCTCAACATCTGTGATACAGGACAGTAACCTTGACAGT 303  
 Qy 81 GYALALAPROHIAEPHOHEUHIATGYTYRGINARGGLNLEUSERTHRYR 100  
 Db 304 GGGTGGCCCCCACCCTTCTGATCGCTACTACAGAGGAGCTGTCCAGACATAC 363  
 Qy 101 ARGAPLEUARGLYSGLYVALTYRVALPROTYRTHNGINGLYYERTRPGIUGIUGLEU 120  
 Db 364 CGGAGCTCCGGAGGGGTGTATGTGCTTACACAGGAGGCAAGGGAGGGAGCTG 423  
 Qy 121 GLYTHASPLEUVALSERILEPROHISGLYPROASNVALTHRVALARGLAASNILEALA 140  
 Db 424 GGACAGCAGCTGTATGACATCCCACTGGCCCCAAGCTACGTGTGGTCCAACTTGCT 483  
 Qy 141 ALAILETHGLUSERASPLYSPHEPHEILEASNGLYSERASNTRPGLUGIYLEUGLY 160  
 Db 484 GCCATCACTGANTCAGACAAAGTTCTTCATCAACAGGCTCCAACTGGAGAGCATCTGGGG 543  
 Qy 161 LEUALATYRALAGIULLEALARGPROASPAPSERLEUGLUPROBHEPHEAPSERLEU 180  
 Db 544 CTGGCTTATGTGATGATGCTGACAGCTCCCTGAGGCTTCTTGTGACTCTCTG 603  
 Qy 181 VALLYEGINTHRHISVALPROASNLEUPHESERLEUGLNUCYSGLYVALAGLYPHEPRO 200  
 Db 604 GTAAAGCAGACCCAGCTCCCAACTCTCTCTGCTGAGCTTGTGTGCTGGCTTCC 663  
 Qy 201 LEUASNGINSEGLIUALLEUALASERVALGYLYSERMETILEILEGLYIIEAP 220  
 Db 664 CTCACCAAGCTGAAGTGTGGCTCTGTGCGAGGAGCATGATCATTTGGAGGTATCGAC 723  
 Qy 221 HISESERLEUTYRTHGLYSERLEUTPYRTHPROLLARGARGIULTPYTYRGLU 240  
 Db 724 CACTCCCTGTACACAGGAGCTCTGTGTATACCACTCGGAGGAGTGTATGTAG 783  
 Qy 241 VALILEGLIETVALRGVALGULLEASNGLYGNASPLEUYSMETASPCYLSYSGIULTYR 260  
 Db 784 GTCATCATGTGTGGGGGAGATCATAGACAGATGTGAAATGTGACTCGAAGAGATAC 843  
 Qy 261 ASNTYRASPLYSERILEVALASPSERGLYTHRTHRASNLEUARGLEUPROLYLYSVAL 280  
 Db 844 AACTATGACAAAGACATGTGGACATGTGGACACCACTTGTGGCCCAAGAAAGT 903  
 Qy 281 PHEGLUALALVALYSESERILEVALIASERSETHRGILYSPHEPROAPGLY 300  
 Db 904 TTGGAAGCTCAGTCAATCATCAAGGACCTCTCCACGAGAAAGTTCCTGTATGT 963  
 Qy 301 PHETRPLEUGLYGLUGLEUVALCYETRPGIUALGLYTHRPROTPAANILEPHE 320  
 Db 964 TTCTGCTAGAGAGAGCTGTGTGTGCTGGCAAGGACCAACCTTGTGAACATTTTC 1023  
 Qy 321 PROVALILESERLEUTYRLEUMETGLYVALTHRASNINSETPHEARGILETHRIE 340  
 Db 1024 CCACTCATCTCACTTCACTTACCTGAGGAGTTACCAACAGTCTTCCCAATCACATC 1083  
 Qy 341 LEUPROGLINGINTYRLEUATGPROVALGIUASPVALAATHISERGINASAPCYSTYR 360  
 Db 1084 CTTCGGAGCAATACCTGGGGCAGGTGGAAGATGTGGCCAGTCCCAAGACGACTGTAC 1143  
 Qy 361 LYEPHEALALIESERGINSESERTHGYTHRVAMETGLYVALAVALIIEMETGIUGLY 380  
 Db 1144 AAGTTGCCATCTCAACAGTATCCAGGGCACTGTATGAGAGCTGTATCATAGAGGCT 1203  
 Qy 381 PHETYRVALVALPHEASPTARGALARGLYSARGILEGLYPHEALVALSERIALCYHIS 400  
 Db 1204 TTCTACCTGTCTTGTATCGGGCCCAAAAGAAATTTGGCTTGTGTCAAGGCTTGGCAT 1263  
 Qy 401 VALHISAPGLUPHEARGTHRALAIALVALGIUGLYPROBHEVALTHRLEUASPMETGLU 420  
 Db 1264 GTGACAGATGATGTTCAAGAGCGGAGGTGGAAGGCTTTTGTCACTTGACATGGAA 1323  
 Qy 421 ASPCYAGLYTYRANILEPROGLINTHRASPGIUSERTHR 433  
 Db 1324 GACTGTGGCTACCAATTCACAGACAGATGATCAACC 1362

RESULT 10  
 AR540886 2070 bp DNA linear PAT 08-OCT-2004  
 LOCUS AR540886  
 DEFINITION Sequence 3 from patent US 6737510.  
 ACCESSION AR540886  
 VERSION AR540886.1 GI:53932399  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 2070)  
 AUTHORS Gurney,M.E., Bienkowski,M.J., Heinrikson,R.L., Parodi,L.A. and Yan,R.  
 TITLE Alzheimer's disease secretase, APP substrates thereof, and uses thereof  
 JOURNAL Patent: US 6737510-A 3 18-MAY-2004;  
 FEATURES Location/Qualifiers  
 source 1..2070  
 /organism="unknown"  
 /mol\_type="genomic DNA"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1,5e-197 Length: 2070  
 Score: 2267.00 Matches: 429  
 Percent Similarity: 99.31% Conservative: 1  
 Best Local Similarity: 99.08% Mismatches: 3  
 Query Match: 99.08% Indels: 0  
 DB: Gaps: 0  
 US-10-726-967a-84 (1-433) x AR540886 (1-2070)  
 Qy 1 THRGIMHISGLYILEARGLEUPROLEUARGSERGLYLEUGLYVALAPROLEUGLYLEU 20  
 Db 64 ACCAGCAGCGGATCCGGCTGCCCTGCGCAGCGGCTGGGGGGGCCCCCTGGGGCTG 123  
 Qy 21 GIULIEASNLEUGLUTHRSPGLUGIUPROGLIUGIUPROGLIARGARGLYSERPHEVAL 40  
 Db 124 CGGCTGCCCCGGGAGACCGACGAGAGAGCCCGAGAGGCCCGGCGGAGGGCAGCTTTGTG 183  
 Qy 41 GIUMETVALASPSNLEUARGGLYLYSERGLYGLINGLYTYRTYRVALGIUMETTHRVAL 60  
 Db 184 GAGATGTGACAACTTGAGGGGCAAGTGGGGCAGGCTTACTACGTGAGATGACCGTG 243  
 Qy 61 GLYSERPROGLINTHRLEUASNILEUVALASPTHRGLYSERSETHRASNHEALVAL 80  
 Db 244 GGACGCCCCCGAGAGCGCTCAACATCTGTGTATACAGGACGACAGTAACTTGTGACAGT 303  
 Qy 81 GYALALAPROHIAEPHOHEUHIATGYTYRGINARGGLNLEUSERTHRYR 100  
 Db 304 GGTGCTGCCCCCACCCTTCTGATCGCTACTACAGAGGAGCTGTCCAGACATAC 363  
 Qy 101 ARGAPLEUARGLYSGLYVALTYRVALPROTYRTHNGINGLYYERTRPGIUGIUGLEU 120  
 Db 364 CGGAGCTCCGGAGGGGTGTATGTGCTTACACAGGAGGCAAGGGAGGGAGCTG 423  
 Qy 121 GLYTHASPLEUVALSERILEPROHISGLYPROASNVALTHRVALARGLAASNILEALA 140  
 Db 424 GGACAGCAGCTGTATGACATCCCACTGGCCCCAAGCTACGTGTGGTCCAACTTGCT 483  
 Qy 141 ALAILETHGLUSERASPLYSPHEPHEILEASNGLYSERASNTRPGLUGIYLEUGLY 160  
 Db 484 GCCATCACTGANTCAGACAAAGTTCTTCATCAACAGGCTCCAACTGGAGCATCTGGGG 543  
 Qy 161 LEUALATYRALAGIULLEALARGPROASPAPSERLEUGLUPROBHEPHEAPSERLEU 180  
 Db 544 CTGGCTTATGTGATGATGCTGACAGCTCCCTGAGGCTTCTTGTGACTCTCTG 603  
 Qy 181 VALLYEGINTHRHISVALPROASNLEUPHESERLEUGLNUCYSGLYVALAGLYPHEPRO 200  
 Db 604 GTAAAGCAGACCCAGCTCCCAACTCTCTCTGCTGAGCTTGTGTGCTGGCTTCC 663

QY 201 LeuAnngInserGIuValleuA1aSerValGIySerMetIleIleGIyIleAsp 220  
Db 664 CTCAACCAAGTGTGAAGTGTGCTCTGTGCGAGAGAGCATGATCGATGAGGATCGAC 723  
QY 221 HisSerLeuTYrThrGIySerLeuTYrThrProIleArgArgIuTYrTYrGIu 240  
Db 724 CACTCGCTGTACACAGGCAAGTCTGTGTATACCACTCCGGGGGAGTGTATTATGAG 783  
QY 241 ValIleIleValArgValGIuIleAnngIyGlnAspLeuIySmetAspCysIyGIuTYr 260  
Db 784 GTCAATCATGTGCGGGGTGAGATCAATGACAGGATGTGAATAATGACTGCAAGAGATAC 843  
QY 261 AsnTYrAspIySerIleValAspSerGIyThrThrAsnLeuArgLeuProIySlyVal 280  
Db 844 AACTATGACAAAGACATTTGTGACAGTGCACCACTTGCTTGCCCAAGAAAGTG 903  
QY 281 PheGIuAlaIleValIySerIleValIleAspIleAspSerThrGIuIySphProAspGIy 300  
Db 904 TTTGAAAGCTCAGTCAAATCCATCAAGGCAAGCTCCCTCCACGAGAAATTCCTGATGGT 963  
QY 301 PheTYrLeuGIyGIuIleuValCysTYrGlnAlaGIyThrThrProTYrAsnIlePhe 320  
Db 964 TTCTGGCTGAGAGAGAGCTGTGTGTGTGCGAAGACAGCACCCCTTGAAACATTTTC 1023  
QY 321 ProValIleSerLeuTYrLeuMetGIyGIuValIleThrAsnGlnSerPheArgIleThrIle 340  
Db 1024 CCAAGTCATCTCACTCACTCACTTATGGGTGAGGTTACCAACCAAGTCCTTCGCAATCACATC 1083  
QY 341 LeuProGIuGlnTYrLeuArgProValGIuAspValAlaIleThrSerGlnAspAspCysTYr 360  
Db 1084 CTTCGCGAGCAATACCTGCGGCGAGTGAAGATGTGGCCACGCTCCMAACAGACTGTTCAC 1143  
QY 361 LysPheAlaIleSerGlnSerSerThrGIyThrValMetGIyAlaValIleMetGIuIy 380  
Db 1144 AAGTTGGCATCTCAACAGTCATCCACGGGCACTGTTATGGAGCTGTTATCATGAGAGGC 1203  
QY 381 PheTYrValIlePheAspArgAlaArgIySargIleGIyPheAlaValSerAlaCysHis 400  
Db 1204 TTCTAGTGTCTTATATCGGGCCGAAAAACGAATTTGCTGTGCTGACGCTTGCCAT 1263  
QY 401 ValHisAspGIuPheArgThrAlaIleValGIuIyProPheValIleThrLeuAspMetGIu 420  
Db 1264 GTTCACGATGATGTTTCAGAGCGGAGCGGTGAAGGCTTTTGTGCACCTTGACATGGA 1323  
QY 421 AspCysGIyTYrAsnIleProGlnThrAspGIuSerThr 433  
Db 1324 GACTGTGCTACACATTCACACAGACAGATGATGATACC 1362

RESULT 11  
ARS60095 2070 bp DNA linear PAT 08-OCT-2004  
LOCUS ARS60095  
DEFINITION Sequence 3 from patent US 6753163.  
ACCESSION ARS60095  
VERSION ARS60095.1 GI:53970462  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 2070)  
AUTHORS Gurney,M.E., Bienkowski,M.J., Heinrichson,R.L., Parodi,L.A. and  
TITLE Alzheimer's disease secretase, APP substrates therefor, and uses  
therefor  
JOURNAL Patent: US 6753163-A 3 22-JUN-2004;  
FEATURES  
source location/Qualifiers  
1..2070  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Alignment Scores: 1..5e-197 Length: 2070  
Pred. No.: 2267.00 Matches: 429  
Score: \*

Percent Similarity: 99.31% Conservative: 1  
Best Local Similarity: 99.08% Mismatches: 3  
Query Match: 99.08% Indels: 0  
DB: 6 Gaps: 0  
US-10-726-967a-84 (1-433) x ARS60095 (1-2070)

QY 1 ThrGlnHisGIyIleArgLeuProLeuArgSerGIyLeuGIyAlaProLeuGIyLeu 20  
Db 64 ACCGAGCAGCGGATCTCGGCTGCCCTGCGCAGCGGCTCGGGGGGCGCCCCCTGGGCTG 123  
QY 21 GluIleAsnLeuGIuThrAspGIuGlnProGIuIuProGIyAArgArgIySerPheVal 40  
Db 124 CGGCTGCCCGGAGAACCCGAGAGAGCCGAGAGAGCCGAGCGGAGGCGAGGCTTTGTG 183  
QY 41 GluMetValAspAsnLeuArgGIyLysSerGIyGlnGIyTYrTYrValGIuMetThrVal 60  
Db 184 GAGATGGTGAACAACCTGAGGGGCAAGTGGGGGCAAGGCTTACATACGAGATGACCTG 243  
QY 61 GlySerProProGIuThrLeuAsnIleLeuValAspThrGIySerSerAsnPheAlaVal 80  
Db 244 GGCAGCCCCCGGAGAGCGCTCAACATCTGTGTGATACAGGCAAGCACTTGTGACATG 303  
QY 81 G1yAlaIaProHisProPheLeuHisArgTYrTYrGlnArgIleuSerSerThrTYr 100  
Db 304 GGTGCTGCCCCCAACCCCTTCTGTGATCGTACTACAGAGGAGAGCTGTCCAGACATAC 363  
QY 101 ArgAspLeuArgIySlyValTYrValProTYrThrGlnIyIySTrGIuGIyGIuLeu 120  
Db 364 CGGAGCTCCCGAAGGGGTGTGTATGTGCTCTACACCAAGGGAAGTGGAAAGGGAGCTG 423  
QY 121 GlyThrAspLeuValSerIleProHisGIyProAsnValIleThrValArgAlaAsnIleAla 140  
Db 424 GGCACCACTCTGTATACATGCCCATGGCCCAAGCTCATGTGTGCTGCAACATTTCT 483  
QY 141 AlaIleThrGlnSerAspLysPhePheIleAsnGIySerAsnTYrGIuIyIleGIy 160  
Db 484 GCCATCACTGATCAGCAAGATTTCTTATCAAGGGCTCCAACTGGAGGACATCTGGGG 543  
QY 161 LeuAlaTYrArgIleAlaArgProAspAspSerLeuGIuProPheAspSerLeu 180  
Db 544 CTGGCTATGTCTAGATTTGCAGGCGCTGACGATCTCCCTGAGGCTTTCTTGTACTCTG 603  
QY 181 ValIySGlnThrHisValProAsnLeuPheSerLeuGlnLeuCysGIyAlaGIyPhePro 200  
Db 604 GTAAAGAGAGACCAAGCTTCCCAACCTTTCTCTGCAAGCTTTGTGTGTGCTTCCCC 663  
QY 201 LeuAnngInserGIuValleuA1aSerValGIyGIySerMetIleIleGIyIleAsp 220  
Db 664 CTCAAACAGCTGAAGTGTGGCTCTGTGCGAGGAGCATGATTCATTTGAGGTATCGAC 723  
QY 221 HisSerLeuTYrThrGIySerLeuTYrTYrThrProIleArgArgIuTYrTYrGIu 240  
Db 724 CACTCGCTGTACACAGGCAAGTCTGTGTATACCACTCCGGGGGAGTGTATTATGAG 783  
QY 241 ValIleIleValArgValGIuIleAnngIyGlnAspLeuIySmetAspCysIyGIuTYr 260  
Db 784 GTCAATCATGTGCGGGGTGAGATCAATGACAGGATCTGAATAATGACTGCAAGAGATAC 843  
QY 261 AsnTYrAspIySerIleValAspSerGIyThrThrAsnLeuArgLeuProIySlyVal 280  
Db 844 AACTATGACAAAGACATTTGTGACAGTGCACACCAACCTTGTTGCCCAAGAAAGTG 903  
QY 281 PheGIuAlaIleValIySerIleValIleAspIleAspSerThrGIuIySphProAspGIy 300  
Db 904 TTTGAAAGCTCAGTCAAATCCATCAAGGCAAGCTCCCTCCACGAGAAATTCCTGATGGT 963  
QY 301 PheTYrLeuGIyGIuIleuValCysTYrGlnAlaGIyThrThrProTYrAsnIlePhe 320  
Db 964 TTCTGGCTGAGAGAGAGCTGTGTGTGTGCGAAGCAGGACCAACCCCTTGGAACATTTTC 1023  
QY 321 ProValIleSerLeuTYrLeuMetGIyGIuValIleThrAsnGlnSerPheArgIleThrIle 340

5  
Db 1024 CCAAGTCATCTCACTTACCTTAATGAGGTACCAACCACTCTTCCGATCAACATC 1083  
Qy 341 LeuProGInGInTYrLeuArgProValGluAspValAlaThrSerGlnAspAspCysTYr 360  
Db 1084 CTTCCGAGCAATATCACTGGCGCAGTGGAGATGTGGCCAGCTCCCAAGACAGACTGTATAC 1143  
Qy 361 LysPheAlaIleSerGlnSerSerThrGlyThrValMetGlyValAlaIleMetGlnGly 380  
Db 1144 AAGTTTGCATCTCAACAGTCACTCCAGCGGCACTGTATATGGAGCTGTATATCATGAGAGGC 1203  
Qy 381 PheTYrValValPheAspArgAlaArgLysArgIleGlyPheAlaValSerAlaCysHis 400  
Db 1204 TTCTAGCTTGTCTTGATGGGCGCCGAAACAAATTGGCTTGTCTGTCAAGCCCTTGCCAT 1263  
Qy 401 ValHisAspGluPheArgThrAlaAlaValGluGlyProPheValThrLeuAspMetGlu 420  
Db 1264 GTGCAGAGATGATTCAGAGCGCAGCGGTGGAGAGGCCCTTTGTCACTTGGACATGGAA 1323  
Qy 421 AspCysGlyTYrAsnIleProGlnThrAspGluSerThr 433  
Db 1324 GACTGTGGCTCAACATTCACAGACAGATGAGTCAACC 1362  
RESULT 12  
AX700447 2070 bp DNA linear PAT 03-APR-2003  
LOCUS AX700447  
DEFINITION Sequence 2 from Patent WO03012089.  
ACCESSION AX700447  
VERSION AX700447.1 GI:29536238  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE 1  
AUTHORS Yon, J., Cleasby, A., Bruinzeel, W.D., Masure, S.L., Tickle, I. and  
Shafit, A.  
TITLE Crystal structure of beta-site app cleaving enzyme (bace) and use  
thereof  
JOURNAL Patent: WO 03012089-A 2 13-FEB-2003;  
Astex Technology Limited (GB) ; JANSSEN PHARMACEUTICA N.V. (BE)  
FEATURES  
source 1..2070  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.5e-197 Length: 2070  
Score: 2267.00 Matches: 429  
Percent Similarity: 99.31% Conservative: 1  
Best Local Similarity: 99.08% Mismatches: 3  
Query Match: 99.08% Indels: 0  
DB: Gaps: 0  
US-10-726-967a-84 (1-433) x AX700447 (1-2070)  
Qy 1 ThrGlnHisGlyValLeuArgLeuProLeuArgSerGlyLeuGlyValAlaProLeuGlyLeu 20  
Db 64 ACCACACAGGAGATCCGGCTGCCCCCTGCGACGCGCTGGGGGCCCCCCCTGGGGGCTG 123  
Qy 21 GluIleAsnLeuGluThrAspGluGluProGluGluProGlyValArgArgIleSerPheVal 40  
Db 124 CGGCTGCCCGGAGAGCCAGAAAGGCCGAGAGACCCGCGGAGGGGCACTTTGTG 183  
Qy 41 GluMetValAspAsnLeuArgGlyLysSerGlyGlnGlyTYrTYrValGluMetThrVal 60  
Db 184 GAGATGTGAGCAACTGAGGGGCAAGTCGGGGCAGGGCTTACTGAGATGAGACCGTG 243  
Qy 61 GlySerProProGlnThrLeuAsnIleLeuValAspThrGlySerSerAsnPheAlaVal 80  
Db 244 GGCAGCCCCCGGAGAGCTCAACATCTGTGGATACAGGCAAGTAACTTTGCAGTG 303

Qy 81 GlyValAlaProHisProPheLeuHisArgTYrTYrGlnArgGlnLeuSerSerThrTYr 100  
Db 304 GATGTGCCCCCAACCCCTTCTGATGCTACTACAGAGGAGAGCTGTCCAGCAACATAC 363  
Qy 101 ArgAspLeuArgLysGlyValTYrValProTYrThrGlnGlyLysTPGlnGlyGluLeu 120  
Db 364 CGGAGCTCCCGAAGGGGTGTATGTGCCCTTACACCCAGGCGCAAGTGGAGGGGAGCTG 423  
Qy 121 GlyThrAspLeuValSerIleProHisGlyProAsnValThrValArgAlaAsnIleAla 140  
Db 424 GGCACCGACCTGGTAAGATCCCCATGGCCCCCAAGCTCATGTGCGGCCAACATCTCT 483  
Qy 141 AlaIleThrGluSerAspLysPhePheIleAsnGlySerAsnTPGlnGlyIleLeuGly 160  
Db 484 GCCATCACTGAATCAGACAAATTTCTTCAATCAACGGCTTCAACTGGAAAGGATCTGGGG 543  
Qy 161 LeuAlaTYrAlaGluIleAlaArgProAspAspSerLeuGluProPhePheAspSerLeu 180  
Db 544 CTGGCTATGTGATGATTCAGAGGCTGACAGACTCCCTGGAGCCCTTCTTGACTCTCTG 603  
Qy 181 ValLysGlnThrHisValProAsnLeuPheSerLeuGlnLeuCysGlyValGlyPhePro 200  
Db 604 GTAAAGCAGACCCAGCTTCCCAACTCTTCTCCGAGCTTTGTGTGCTGCTGCC 663  
Qy 201 LeuAsnGlnSerGluValLeuAlaSerValGlyGlySerMetIleIleGlyValIleAsp 220  
Db 664 CTCAACAGCTTGAAGTCTGAGCTCTCTTCCGAGGAGCAATCATATTGGAGATTCAC 723  
Qy 221 HisSerLeuTYrThrGlySerLeuTPYrTYrThrProIleArgArgGluTPYrTYrGlu 240  
Db 724 CACTGCTGTACAGAGCAGCTCTGTATACACCATCCGCGGAGATGATATATAG 783  
Qy 241 ValIleIleValArgValGluIleAsnGlyLysAspLeuLysMetAspCysGlyGlyTYr 260  
Db 784 GTCAATCATTTGGCGGTGTGAATCATATGACAGATTCGAAATGAGTGCAGGAAGATAC 843  
Qy 261 AsnTYrAspLysSerIleValAspSerGlyThrThrAsnLeuArgLeuProLysVal 280  
Db 844 AACTATGACAAAGCATTTGTGACATGTGGACATGGCACACCAACTTCTTGGCCAGAAAGTG 903  
Qy 281 PheGluAlaAlaValLysSerIleLysAlaIleAspSerThrGluLysPheProAspGly 300  
Db 904 TTTGAGAGCTCAGATCAATTCATCAAGGACGCTCTCAAGAGAAATTCCTGATGAT 963  
Qy 301 PheTPLeuGlyGlnGlnLeuValCysTPGlnAlaGlyThrThrProTPAsnIlePhe 320  
Db 964 TTCTGGCTAGGAGACAGCTGTGTGTGGCAAGCAGCACCACTTGGAAACATTTTC 1023  
Qy 321 ProValIleSerLeuTYrLeuMetGlyGluValThrAsnGlnSerPheArgIleThrIle 340  
Db 1024 CCAAGTCATCTCACTTACCTTAATGAGGTAGATTACCAACCAAGTCTTCCGATCAACATC 1083  
Qy 341 LeuProGInGInTYrLeuArgProValGluAspValAlaThrSerGlnAspAspCysTYr 360  
Db 1084 CTTCCGAGCAATATCACTCGGCACTGGAAGATGTGGCAGGTCCCAAGACAGACTGTATC 1143  
Qy 361 LysPheAlaIleSerGlnSerSerThrGlyThrValMetGlyValAlaIleMetGlnGly 380  
Db 1144 AAGTTTGCATCTCAACAGTCACTCCAGCGGCACTGTATATGGAGCTGTATATCATGAGAGGC 1203  
Qy 381 PheTYrValValPheAspArgAlaArgLysArgIleGlyPheAlaValSerAlaCysHis 400  
Db 1204 TTCTAGCTTGTCTTGATGGGCGCCGAAACAAATTGGCTTGTCTGTCAAGCCCTTGCCAT 1263  
Qy 401 ValHisAspGluPheArgThrAlaAlaValGluGlyProPheValThrLeuAspMetGlu 420  
Db 1264 GTGCAGAGATGATTCAGAGCGCAGCGGTGGAGAGGCCCTTTGTCACTTGGACATGGAA 1323  
Qy 421 AspCysGlyTYrAsnIleProGlnThrAspGluSerThr 433  
Db 1324 GACTGTGGCTCAACATTCACAGACAGATGAGTCAACC 1362  
RESULT 13



AF200343 2070 bp mRNA linear PRI 12-DEC-1999  
LOCUS Homo sapiens chromosome 11 aspartyl protease 2 mRNA, complete cds.  
DEFINITION AF200343  
ACCESSION AF200343.1 GI:6561813  
VERSION AF200343.1  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
AUTHORS Yan, R., Bienkowsk, M.J., Shuck, M.E., Mao, H., Tor, M.C.,  
Pauley, A.M., Braehier, J.R., Stracman, N.C., Mathews, W.R., Buhl, A.E.,  
Carter, D.B., Tomasselli, A.G., Parodi, L.A., Henrikson, R.L. and  
Gurney, M.E.  
TITLE Membrane-anchored aspartyl protease with Alzheimer's disease  
beta-secretase activity  
JOURNAL Nature 402 (6761), 533-537 (1999)  
MEDLINE 20057170  
PUBMED 10591213  
REFERENCE 2 (bases 1 to 2070)  
AUTHORS Bienkowsk, M.J., Shuck, M.E., Slightom, J.L. and Dong, R.F.  
TITLE Direct Submission  
JOURNAL Submitted (29-OCT-1999) Genomics Research, Pharmacidrupjohn, 301  
Henrietta, Kalamazoo, MI 49007, USA  
FEATURES  
source  
1..2070  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="11"  
/map="11q23.3-24.1"  
1..1506  
/codon\_start=1  
/product="aspartyl protease 2"  
/protein\_id="AAPI7079.1"  
/db\_xref="GI:6561814"  
/translation="MAQALFWLLIMWAGAVLPAAHGTQHGIRLPRLRSGLGAGAPLRLP  
REYDEEPEERGRGSEVENVDMRGSGGGYVEMTVGSPPTLILVDTGSSNFAVG  
AAPHPLHRYOQLSTYADLAKGVYVADQKMEGELGTLSVHPGPNYVANI  
AALTSDKFFINGSNMEGILGLAYAIAPDLSLEFPRLVYKOTHHVNPICLGA  
GFLNQSSEVLASVSGSMIGGIDHSLYTSGLVTPRRMYEVYIVAYEINGDLKX  
DCKEYVDSIVDSGTTNRLPKYVEAAVKSIIKASSTREKPDGFMLGSDLVQAG  
TTPWNI PPIVSLYLNKESTVNSFRITLIPQYLRVEDVAISQDDCYFAISQSTGT  
VNGAVVMEPTVYFEDARRIRIGFAVSAACHDEFTPAVEGPTVLDMEGYNIPOT  
DSTLTMTIAYVMAAICALFMLPLCLMVCOMRCIRLRQOHDFADISLX"

ORIGIN  
Alignment Scores:  
Pred. No.: 1.5e-197 Length: 2070  
Score: 2267.00 Matches: 429  
Percent Similarity: 99.318 Conservative: 1  
Best Local Similarity: 99.084 Mismatches: 3  
Query Match: 99.084 Indels: 0  
Gaps: 0  
US-10-726-967a-84 (1-433) x AF200343 (1-2070)

QY 1 ThrGlnHISGLYIleArgLeuProLeuArgSerGlyLeuGlyGlyAlaProLeuGlyLeu 20  
Db 64 ACCGACGACGGGATCCGGCTGCCCTGGCGACGGGCTGGGGGGGCCCCCTGGGGCTG 123  
QY 21 GtUleAenLeuGtThrAspGlnGluProGlnGluProGlyAlaGlyGlySerPheVal 40  
Db 124 CGGCTGCCCCCGGAGAACGACGAGAGGCCCGGCGCGAGGGGCGACCTTTGTG 183  
QY 41 GtUmetValAspAsnLeuArgGlyLysSerGlyGlnGlyTyTyTyValGtUmetThrVal 60  
Db 184 GAGATGTGTGACCACTGAGGGGCAAGTCGGGGCGACGGCTTCTACGTGGAGATACCGTG 243  
QY 61 GlySerProProGlnInThLeuAsnIleUeuValAspThrgLysSerAspAspAlaVal 80  
Db 244 GCGAGCCCCCGGACGCTCAACATCTCTGTGTGATACAGGACGACGACGACGACGACG 303

QY 81 GlyAlaAlaProHisProPheLeuHISArgTyTyTyGlnArgGlnLeuSerSerThrTy 100  
Db 304 GGTGCTGCCCCCGACCCCTTCTCTGATGCTACCTACACGAGAGGAGCTGTCCAGACATAC 363  
QY 101 ArgAspLeuArgLysGlyValTyTyValProTyThrgInGlyLysTyPrGlnGlyLeu 120  
Db 364 CGGACCTCCCGGAGAGGTGTGTATGTGCTTACACCCGAGGCGAAGTGGAGAGGAGACTG 423  
QY 121 GlyThrAspLeuValSerIleProHisGlyProAsnValThrValAlaAsnIleAla 140  
Db 424 GGCACCGACCTGGTGAAGATCCCGGACGACGACGACGACGACGACGACGACGACGACG 483  
QY 141 AlaIleThrgLysSerAspLysPheIleAsnGlySerAsnTyPrGlnGlyIleLeuGly 160  
Db 484 GCCATCATGTGATGACCAAGTCTTATCATCAACGGCTTCAACTGGAGAGGACATCTGGGG 543  
QY 161 LeuAlaTyTyAlaGlnIleAlaArgProAspAspSerLeuGluProPheAspSerLeu 180  
Db 544 CTGGCTATGTGTGAGATGTGCGAGGCTGAGGACCTCCCTGGAGGCTTCTTGTACTCTG 603  
QY 181 ValLysGlnThrHisValProAsnLeuPheSerLeuGlnLeuGlyAlaGlyPhePro 200  
Db 604 GTAAAGACGACCCACGTTCCACCTTCTTCCCTGACGCTTGTGTGCTGCTGCTTCC 663  
QY 201 LeuAsnGlnSerGlnValIleuAlaSerValGlyGlySerMetIleIleGlyIleAsp 220  
Db 664 CTCAACGATCTGAGATGTGCTGCTGTGCGAGGAGCATGATCATTTGGAGGTATCGAC 723  
QY 221 HisSerLeuTyTyThrgLysSerLeuTyTyThrProIleArgArgGluTyTyTyGlu 240  
Db 724 CACTCGCTGTACAGGACAGTCTGTGTATACACCATCCGGGGGAGTGTATATAGAG 783  
QY 241 ValIleIleValArgValGluIleAsnGlyGlnAspLeuLysMetAspCysLysGlyTy 260  
Db 784 GTCATCATTTGTGGGGGAGATCAATGACAGGATCTGAATGACATGCAAGAGATGAC 843  
QY 261 AsnTyTyAspLysSerIleValAspSerGlyThrThrAsnLeuArgLeuProLysVal 280  
Db 844 AACTATGACAAAGCATGTGTGACAGTGGACACCAACCATCTGTTGCGCAAGAAAGTG 903  
QY 281 PheGlnAlaAlaValLysSerIleLysAlaAsnSerThrGlnLysPheProAspGly 300  
Db 904 TTGAAGCTGCAATCAATCAATCAAGGACCTCTCCAGGAGAGATCCCTGATGGT 963  
QY 301 PheTrpLeuGlyGlnGlnLeuValCysTyPrGlnAlaGlyThrThrProTrpAsnIlePhe 320  
Db 964 TTCTGGCTAGAGAGAGAGCTGTGTGCTGGCAAGGACGACACCCCTTGGAAATTTTC 1023  
QY 321 ProValIleSerLeuTyTyLeuMetGlyGlyValIleThrgInSerPheArgIleThrIle 340  
Db 1024 CGAGTATCTCACTCACTCAATGAGTGTGATGATGATGATGATGATGATGATGATGATG 1083  
QY 341 LeuProGlnGlnTyTyLeuArgProValGlnAspValAlaIleThrgInSerGlnAspAspCysTy 360  
Db 1084 CTTCCGACCAATATACCTGGCGGACGAGTGAAGTGTGGCAGCTCCCAAGAGACTGTAC 1143  
QY 361 LysPheAlaIleSerGlnSerSerThrgTyThrValMetGlyAlaValIleMetGlnGly 380  
Db 1144 AAGTTTGCCATCTCAAGATCAACGAGCTGTATGAGGAGCTGTATGATGAGAGGAGC 1203  
QY 381 PheTyTyValValPheAspArgAlaArgLysArgIleGlyPheAlaValIleSerAlaCysHis 400  
Db 1204 TTCTAGTTGTCTTGTGATCGGGCCGGAABAAGATTTGGCTTGTCTGACGAGCTTGGCAT 1263  
QY 401 ValHisAspGluPheArgThrAlaAlaValGlnGlyProPheValThrLeuAspMetGlu 420  
Db 1264 GTGACGATGAGTGTCAAGACGCGAGCGGTGGAAGGCTTTTGTGCACTTGACATGGAA 1323  
QY 421 AspCysGlyTyTyAsnIleProGlnThrAspGlnSerThr 433  
Db 1324 GACTGTGCTACCAACATTTCCACAGACGATGATGATGATGATGATGATGATGATGATG 1362







GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using SW model

Run on: August 22, 2005, 00:19:29 ; Search time 1527 Seconds  
(without alignments)  
111.039 Million cell updates/sec

Title: US-10-726-967A-84

Perfect score: 2288  
Sequence: 1 TQHGIRLPRLRSGIGAPLGL.....FVTLMDGCGINIPQTDST 433

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1759131 seqs, 391586102 residues

Total number of hits satisfying chosen parameters: 1759131

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA.\*  
1: /cgn2\_6/ptodata/1/pubppaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubppaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubppaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubppaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubppaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubppaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubppaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubppaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubppaa/US09\_NEW\_PUB.pep.\*  
12: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/1/pubppaa/US10\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubppaa/US11\_PUBCOMB.pep.\*  
20: /cgn2\_6/ptodata/1/pubppaa/US11\_PUBCOMB.pep.\*  
21: /cgn2\_6/ptodata/1/pubppaa/US60\_NEW\_PUB.pep.\*  
22: /cgn2\_6/ptodata/1/pubppaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2288	100.0	433	17	US-10-726-967A-84
2	2286	99.9	433	17	US-10-726-967A-81
3	2273	99.3	433	17	US-10-726-967A-78
4	2267	99.1	435	15	US-10-400-273-2
5	2267	99.1	456	16	US-10-281-092-8
6	2267	99.1	480	20	US-11-089-918-66
7	2267	99.1	480	20	US-11-090-866-66
8	2267	99.1	480	20	US-11-069-377-66
9	2267	99.1	480	20	US-11-090-872-66
10	2267	99.1	488	9	US-09-796-264-2
11	2267	99.1	488	9	US-09-845-226-2

12	2267	99.1	488	9	US-09-795-903A-2	Sequence 2, Appl1
13	2267	99.1	488	14	US-10-032-818-2	Sequence 2, Appl1
14	2267	99.1	488	16	US-10-820-853-2	Sequence 2, Appl1
15	2267	99.1	488	16	US-10-773-754-2	Sequence 2, Appl1
16	2267	99.1	501	14	US-10-032-818-4	Sequence 4, Appl1
17	2267	99.1	501	14	US-10-214-932-104	Sequence 104, App
18	2267	99.1	501	15	US-10-372-473-1	Sequence 1, Appl1
19	2267	99.1	501	15	US-10-652-927-4	Sequence 4, Appl1
20	2267	99.1	501	15	US-10-652-830-4	Sequence 4, Appl1
21	2267	99.1	501	16	US-10-281-092-6	Sequence 6, Appl1
22	2267	99.1	501	16	US-10-466-258-2	Sequence 2, Appl1
23	2267	99.1	501	16	US-10-652-045-4	Sequence 4, Appl1
24	2267	99.1	501	16	US-10-343-389A-19	Sequence 285, App
25	2267	99.1	501	16	US-10-723-860-285	Sequence 1, Appl1
26	2267	99.1	501	16	US-10-837-021A-1	Sequence 1, Appl1
27	2267	99.1	501	17	US-10-940-867-4	Sequence 4, Appl1
28	2267	99.1	501	17	US-10-726-967A-1	Sequence 1, Appl1
29	2267	99.1	501	18	US-10-466-391A-2	Sequence 2, Appl1
30	2267	99.1	501	20	US-11-089-918-2	Sequence 2, Appl1
31	2267	99.1	501	20	US-11-090-866-2	Sequence 2, Appl1
32	2267	99.1	501	20	US-11-059-377-2	Sequence 2, Appl1
33	2267	99.1	501	20	US-11-090-867-3	Sequence 3, Appl1
34	2267	99.1	503	9	US-09-796-264-3	Sequence 3, Appl1
35	2267	99.1	503	9	US-09-845-226-3	Sequence 3, Appl1
36	2267	99.1	503	9	US-09-795-903A-3	Sequence 3, Appl1
37	2267	99.1	503	14	US-10-032-818-3	Sequence 3, Appl1
38	2267	99.1	503	16	US-10-820-953-3	Sequence 3, Appl1
39	2267	99.1	503	16	US-10-773-754-3	Sequence 3, Appl1
40	2267	99.1	509	16	US-10-275-339A-7	Sequence 7, Appl1
41	2262	98.9	453	9	US-09-794-927-30	Sequence 30, Appl1
42	2262	98.9	453	9	US-09-795-847-30	Sequence 30, Appl1
43	2262	98.9	453	9	US-09-794-743-30	Sequence 30, Appl1
44	2262	98.9	453	9	US-09-794-746-30	Sequence 30, Appl1
45	2262	98.9	453	9	US-09-794-925-30	Sequence 30, Appl1

## ALIGNMENTS

RESULT 1  
US-10-726-967A-84  
Sequence 84, Application US/10726967A  
Publication No. US20050074456A1  
GENERAL INFORMATION:  
APPLICANT: Ballinger, Marcus  
TITLE OF INVENTION: Constructs for Homogenously Processed Preparations of Beta Site  
FILE REFERENCE: 2004345-0021  
CURRENT APPLICATION NUMBER: US/10/726,967A  
CURRENT FILING DATE: 2003-12-02  
NUMBER OF SEQ ID NOS: 110  
SOFTWARE: Patentin version 3.2  
SEQ ID NO 84  
LENGTH: 433  
TYPE: PRT  
ORGANISM: Artificial  
FEATURE:  
OTHER INFORMATION: Soluble human PROBACE1 containing an engineered E1NL site  
US-10-726-967A-84

Query Match 100.0%; Score 2288; DB 17; Length 433;  
Best Local Similarity 100.0%; Pred. No. 4.5e-211; Indels 0; Gaps 0;  
Matches 433; Conservative 0; Mismatches 0;  
QY 1 TQHGIRLPRLRSGIGAPLGLBINLETDEPERGRGSEFVEMVDNLRGSGGGYVEMTV 60  
DB 1 TQHGIRLPRLRSGIGAPLGLBINLETDEPERGRGSEFVEMVDNLRGSGGGYVEMTV 60  
QY 61 GSPPTLNLIVDTGSSNFAVGAAPHPFLHRYRQRLSSYTRDLRKGVVYPYQKWEGL 120  
DB 61 GSPPTLNLIVDTGSSNFAVGAAPHPFLHRYRQRLSSYTRDLRKGVVYPYQKWEGL 120

```
Qy 121 GTDLVSIHPGPNVTVRANIAATTESDKFFINGSNMEGILGLAYAEIARPDLSLEPPFDSL 180
| | | | |
Db 121 GTDLVSIHPGPNVTVRANIAATTESDKFFINGSNMEGILGLAYAEIARPDLSLEPPFDSL 180
Qy 181 VKQTHVNLFSIQLCGAGFPINQSEVLASVGGSMIIIGIDHSLYTGSLMYTPIRREMYE 240
| | | | |
Db 181 VKQTHVNLFSIQLCGAGFPINQSEVLASVGGSMIIIGIDHSLYTGSLMYTPIRREMYE 240
Qy 241 VIVRVEINGODLKMDCKEYNVDKSIYDSGTTNLRPKKVEAAVKSIIKAASSTKEPFDG 300
| | | | |
Db 241 VIVRVEINGODLKMDCKEYNVDKSIYDSGTTNLRPKKVEAAVKSIIKAASSTKEPFDG 300
Qy 301 FMLEGQLVCMQAGTTPMNIFFVVISLYLMGEVYNOSFRITTLIPQOYLRAVEDVATSODDCY 360
| | | | |
Db 301 FMLEGQLVCMQAGTTPMNIFFVVISLYLMGEVYNOSFRITTLIPQOYLRAVEDVATSODDCY 360
Qy 361 KFAISQSSGTVMGAVIMEGFYVVDRAKRIIGFAVSACHVDEFRTAAVEGPFVTLTME 420
| | | | |
Db 361 KFAISQSSGTVMGAVIMEGFYVVDRAKRIIGFAVSACHVDEFRTAAVEGPFVTLTME 420
Qy 421 DCGYNIPQTDTEST 433
| | | | |
Db 421 DCGYNIPQTDTEST 433
```

## RESULT 2

```
US-10-726-967a-81
; Sequence 81, Application US/10726967a
; Publication No. US20050074456A1
; GENERAL INFORMATION:
; APPLICANT: Ballinger, Marcus
; TITLE OF INVENTION: Constructs for Homogenously Processed Preparations of Beta Site
; FILE REFERENCE: 2004345-0021
; CURRENT APPLICATION NUMBER: US/10/726,967a
; CURRENT FILING DATE: 2003-12-02
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 81
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Soluble human proBACE1
US-10-726-967a-81
```

```
Query Match 99.9%; Score 2286; DB 17; Length 433;
Best Local Similarity 99.8%; Pred. No. 7e-211;
Matches 432; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 TQHGIRLPLRSGLGAPLGLLEINLETDEBEPERGRGSFVEMVDNLRGKSGGYVEMTV 60
| | | | |
Db 1 TQHGIRLPLRSGLGAPLGLLEINLETDEBEPERGRGSFVEMVDNLRGKSGGYVEMTV 60
Qy 61 GSPPQTLNIIIVDTGSSNFAVGAAPHPFLHRYYQROLSTYRDLRKGVVPTYGKMEGEL 120
| | | | |
Db 61 GSPPQTLNIIIVDTGSSNFAVGAAPHPFLHRYYQROLSTYRDLRKGVVPTYGKMEGEL 120
Qy 121 GTDLVSIHPGPNVTVRANIAATTESDKFFINGSNMEGILGLAYAEIARPDLSLEPPFDSL 180
| | | | |
Db 121 GTDLVSIHPGPNVTVRANIAATTESDKFFINGSNMEGILGLAYAEIARPDLSLEPPFDSL 180
Qy 181 VKQTHVNLFSIQLCGAGFPINQSEVLASVGGSMIIIGIDHSLYTGSLMYTPIRREMYE 240
| | | | |
Db 181 VKQTHVNLFSIQLCGAGFPINQSEVLASVGGSMIIIGIDHSLYTGSLMYTPIRREMYE 240
Qy 241 VIVRVEINGODLKMDCKEYNVDKSIYDSGTTNLRPKKVEAAVKSIIKAASSTKEPFDG 300
| | | | |
Db 241 VIVRVEINGODLKMDCKEYNVDKSIYDSGTTNLRPKKVEAAVKSIIKAASSTKEPFDG 300
Qy 301 FMLEGQLVCMQAGTTPMNIFFVVISLYLMGEVYNOSFRITTLIPQOYLRAVEDVATSODDCY 360
| | | | |
Db 301 FMLEGQLVCMQAGTTPMNIFFVVISLYLMGEVYNOSFRITTLIPQOYLRAVEDVATSODDCY 360
```

```
Qy 361 KFAISQSSGTVMGAVIMEGFYVVDRAKRIIGFAVSACHVDEFRTAAVEGPFVTLTME 420
| | | | |
Db 361 KFAISQSSGTVMGAVIMEGFYVVDRAKRIIGFAVSACHVDEFRTAAVEGPFVTLTME 420
Qy 421 DCGYNIPQTDTEST 433
| | | | |
Db 421 DCGYNIPQTDTEST 433
```

## RESULT 3

```
US-10-726-967a-78
; Sequence 78, Application US/10726967a
; Publication No. US20050074456A1
; GENERAL INFORMATION:
; APPLICANT: Ballinger, Marcus
; TITLE OF INVENTION: Constructs for Homogenously Processed Preparations of Beta Site
; FILE REFERENCE: 2004345-0021
; CURRENT APPLICATION NUMBER: US/10/726,967a
; CURRENT FILING DATE: 2003-12-02
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 78
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Soluble human proBACE1 containing an engineered RUPL site
US-10-726-967a-78
```

```
Query Match 99.3%; Score 2273; DB 17; Length 433;
Best Local Similarity 99.3%; Pred. No. 1.2e-209;
Matches 430; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy 1 TQHGIRLPLRSGLGAPLGLLEINLETDEBEPERGRGSFVEMVDNLRGKSGGYVEMTV 60
| | | | |
Db 1 TQHGIRLPLRSGLGAPLGLLEINLETDEBEPERGRGSFVEMVDNLRGKSGGYVEMTV 60
Qy 61 GSPPQTLNIIIVDTGSSNFAVGAAPHPFLHRYYQROLSTYRDLRKGVVPTYGKMEGEL 120
| | | | |
Db 61 GSPPQTLNIIIVDTGSSNFAVGAAPHPFLHRYYQROLSTYRDLRKGVVPTYGKMEGEL 120
Qy 121 GTDLVSIHPGPNVTVRANIAATTESDKFFINGSNMEGILGLAYAEIARPDLSLEPPFDSL 180
| | | | |
Db 121 GTDLVSIHPGPNVTVRANIAATTESDKFFINGSNMEGILGLAYAEIARPDLSLEPPFDSL 180
Qy 181 VKQTHVNLFSIQLCGAGFPINQSEVLASVGGSMIIIGIDHSLYTGSLMYTPIRREMYE 240
| | | | |
Db 181 VKQTHVNLFSIQLCGAGFPINQSEVLASVGGSMIIIGIDHSLYTGSLMYTPIRREMYE 240
Qy 241 VIVRVEINGODLKMDCKEYNVDKSIYDSGTTNLRPKKVEAAVKSIIKAASSTKEPFDG 300
| | | | |
Db 241 VIVRVEINGODLKMDCKEYNVDKSIYDSGTTNLRPKKVEAAVKSIIKAASSTKEPFDG 300
Qy 301 FMLEGQLVCMQAGTTPMNIFFVVISLYLMGEVYNOSFRITTLIPQOYLRAVEDVATSODDCY 360
| | | | |
Db 301 FMLEGQLVCMQAGTTPMNIFFVVISLYLMGEVYNOSFRITTLIPQOYLRAVEDVATSODDCY 360
Qy 361 KFAISQSSGTVMGAVIMEGFYVVDRAKRIIGFAVSACHVDEFRTAAVEGPFVTLTME 420
| | | | |
Db 361 KFAISQSSGTVMGAVIMEGFYVVDRAKRIIGFAVSACHVDEFRTAAVEGPFVTLTME 420
Qy 421 DCGYNIPQTDTEST 433
| | | | |
Db 421 DCGYNIPQTDTEST 433
```

## RESULT 4

```
US-10-400-273-2
; Sequence 2, Application US/10400273
; Publication No. US20040014194A1
```

```

; GENERAL INFORMATION:
; APPLICANT: Beyer, Brian
; APPLICANT: Hammond, Gerald S
; APPLICANT: Reichert, Paul
; APPLICANT: Strickland, Corey
; APPLICANT: Wang, Wenyan
; APPLICANT: Weber, Patricia C
; APPLICANT: Wong, Gwendolyn
; APPLICANT: Zhang, Lili
; TITLE OF INVENTION: BETA-SECRETASE CRYSTALS AND METHODS FOR PREPARING AND USING THE S
; FILE REFERENCE: J801531-K-US
; CURRENT APPLICATION NUMBER: US/10/400,273
; CURRENT FILING DATE: 2003-03-26
; PRIOR APPLICATION NUMBER: 60/367,937
; PRIOR FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 435
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-400-273-2

Query Match      99.1%; Score 2267; DB 15; Length 435;
Best Local Similarity 99.1%; Pred. No. 4.7e-209;
Matches 429; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 TQHGRLPLRSGLGAPLGLINLETDEBEPGRGRSGFVEMVDNLRGSGGGYVEMTV 60
DB 1 TQHGRLPLRSGLGAPLGLINLETDEBEPGRGRSGFVEMVDNLRGSGGGYVEMTV 60
QY 61 GSPPTNLILVDTGSSNFAVGAAHPHPLHRYRQRLSTYRDLRKGVVYPYQGWEGEL 120
DB 61 GSPPTNLILVDTGSSNFAVGAAHPHPLHRYRQRLSTYRDLRKGVVYPYQGWEGEL 120
QY 121 GTDLVSIPIHGPVNTYRANIAATTESDKPFINSNMGGILGLAETARPDSDLBPFDL 180
DB 121 GTDLVSIPIHGPVNTYRANIAATTESDKPFINSNMGGILGLAETARPDSDLBPFDL 180
QY 181 VKQTHVFNLFSLQLCGAGFPLNQSEVLAASVSGSMIIIGIDHSLYTSLWYTPIRREMYE 240
DB 181 VKQTHVFNLFSLQLCGAGFPLNQSEVLAASVSGSMIIIGIDHSLYTSLWYTPIRREMYE 240
QY 241 VIIIRVEINGODLKMDCKEYNDKSIYDSGTTNLRPKVFEAAVKSIAASTTEKFPDG 300
DB 241 VIIIRVEINGODLKMDCKEYNDKSIYDSGTTNLRPKVFEAAVKSIAASTTEKFPDG 300
QY 301 FWLGEOLVCMQAGTTTWNIFPVISLYLMGEVNTQSFRTITLPOQYLRPEVDVATSDDDCY 360
DB 301 FWLGEOLVCMQAGTTTWNIFPVISLYLMGEVNTQSFRTITLPOQYLRPEVDVATSDDDCY 360
QY 361 KPFISSSTGTVMGAVIMGFFVVPDRARKRIGFAVSACHVHDEFRTAAVEGPFVTLDM 420
DB 361 KPFISSSTGTVMGAVIMGFFVVPDRARKRIGFAVSACHVHDEFRTAAVEGPFVTLDM 420
QY 421 DCGYNIPQTDST 433
DB 421 DCGYNIPQTDST 433

RESULT 5
US-10-281-092-8
; Sequence 8, Application US/10281092
; Publication No. US20040121947A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Arun K.
; APPLICANT: Tang, Jordan J.N.
; APPLICANT: Bilcer, Geoffrey
; APPLICANT: Chang, Wenpin
; APPLICANT: Hong, Lin
; APPLICANT: Koelsch, Gerald E.
; APPLICANT: Loy, Jeffrey A.
; APPLICANT: Turner, Robert T., III

```

```

; APPLICANT: Devasumadrum, Thippeswamy
; TITLE OF INVENTION: COMPOUNDS WHICH INHIBIT BETA-SECRETASE
; FILE REFERENCE: 2952.1001-004
; CURRENT APPLICATION NUMBER: US/10/281,092
; CURRENT FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: US 10/032,818
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: PCT US01/50826
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: US 60/258,705
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: US 60/275,756
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: US 60/335,952
; PRIOR FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: US 60/333,545
; PRIOR FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: US 60/348,464
; PRIOR FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/348,615
; PRIOR FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/390,804
; PRIOR FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: US 60/397,557
; PRIOR FILING DATE: 2002-07-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FaalSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 456
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: deduced amino acid sequence of promemapsin 2-T1
; US-10-281-092-8

Query Match      99.1%; Score 2267; DB 16; Length 456;
Best Local Similarity 99.1%; Pred. No. 5.1e-209;
Matches 429; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 TQHGRLPLRSGLGAPLGLINLETDEBEPGRGRSGFVEMVDNLRGSGGGYVEMTV 60
DB 24 TQHGRLPLRSGLGAPLGLINLETDEBEPGRGRSGFVEMVDNLRGSGGGYVEMTV 83
QY 61 GSPPTNLILVDTGSSNFAVGAAHPHPLHRYRQRLSTYRDLRKGVVYPYQGWEGEL 120
DB 84 GSPPTNLILVDTGSSNFAVGAAHPHPLHRYRQRLSTYRDLRKGVVYPYQGWEGEL 143
QY 121 GTDLVSIPIHGPVNTYRANIAATTESDKPFINSNMGGILGLAETARPDSDLBPFDL 180
DB 144 GTDLVSIPIHGPVNTYRANIAATTESDKPFINSNMGGILGLAETARPDSDLBPFDL 203
QY 181 VKQTHVFNLFSLQLCGAGFPLNQSEVLAASVSGSMIIIGIDHSLYTSLWYTPIRREMYE 240
DB 204 VKQTHVFNLFSLQLCGAGFPLNQSEVLAASVSGSMIIIGIDHSLYTSLWYTPIRREMYE 263
QY 241 VIIIRVEINGODLKMDCKEYNDKSIYDSGTTNLRPKVFEAAVKSIAASTTEKFPDG 300
DB 264 VIIIRVEINGODLKMDCKEYNDKSIYDSGTTNLRPKVFEAAVKSIAASTTEKFPDG 323
QY 301 FWLGEOLVCMQAGTTTWNIFPVISLYLMGEVNTQSFRTITLPOQYLRPEVDVATSDDDCY 360
DB 324 FWLGEOLVCMQAGTTTWNIFPVISLYLMGEVNTQSFRTITLPOQYLRPEVDVATSDDDCY 383
QY 361 KPFISSSTGTVMGAVIMGFFVVPDRARKRIGFAVSACHVHDEFRTAAVEGPFVTLDM 420
DB 384 KPFISSSTGTVMGAVIMGFFVVPDRARKRIGFAVSACHVHDEFRTAAVEGPFVTLDM 443
QY 421 DCGYNIPQTDST 433
DB 444 DCGYNIPQTDST 456

```

```
RESULT 6
US-11-089-918-66
/ Sequence 66, Application US/11089918
/ Publication No. US20050164327A1
/ GENERAL INFORMATION:
/ APPLICANT: Anderson, John P.
/ APPLICANT: Basi, Gurigbal
/ APPLICANT: Doane, Minh Tam
/ APPLICANT: Frigon, Normand
/ APPLICANT: John, Varghese
/ APPLICANT: Power, Michael
/ APPLICANT: Sinha, Sukanto
/ APPLICANT: Tatsuno, Gwen
/ APPLICANT: Tung, Jay
/ APPLICANT: Wang, Shuwen
/ APPLICANT: McConlogue, Lisa
/ TITLE OF INVENTION: Beta-Secretase Enzyme Compositions and
/ FILE REFERENCE: 228-US-NEW2C9
/ CURRENT FILING DATE: 2005-03-25
/ PRIOR APPLICATION NUMBER: US/11/089,918
/ PRIOR FILING DATE: 2000-11-28
/ PRIOR APPLICATION NUMBER: US 09/501,708
/ PRIOR FILING DATE: 2000-02-10
/ PRIOR APPLICATION NUMBER: US 09/471,669
/ PRIOR FILING DATE: 1999-12-24
/ PRIOR APPLICATION NUMBER: 60/119,571
/ PRIOR FILING DATE: 1999-02-10
/ PRIOR APPLICATION NUMBER: 60/139,172
/ PRIOR FILING DATE: 1999-06-15
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 66
/ LENGTH: 480
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-11-089-918-66

Query Match      99.1%; Score 2267; DB 20; Length 480;
Best Local Similarity 99.1%; Pred. No. 5.5e-209;
Matches 429; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      1 TOHGIRLPRLSGGAPLGLIEINLETDEPEEPGRGSGFVEMVDNLRGSGGGYVEMTV 60
DB      1 TOHGIRLPRLSGGAPLGLRLPRETDEPEEPGRGSGFVEMVDNLRGSGGGYVEMTV 60
QY      61 GSPPTLNLIVDTGSSNFVGAAPHPFLHRYYQROLSSTYRDLRGVYVPYTOGMEBEL 120
DB      61 GSPPTLNLIVDTGSSNFVGAAPHPFLHRYYQROLSSTYRDLRGVYVPYTOGMEBEL 120
QY      121 GTDLVSIPIHGPVNTYRANIAITTESDKPFINSNMEGLIGLAVAIAPDDSLPEFPSTL 180
DB      121 GTDLVSIPIHGPVNTYRANIAITTESDKPFINSNMEGLIGLAVAIAPDDSLPEFPSTL 180
QY      181 VQOTHPNLFSLQCGAGFPLNQSEVLASVGSMTIGSIDLSLYTGSIMTYPIRREWEYE 240
DB      181 VQOTHPNLFSLQCGAGFPLNQSEVLASVGSMTIGSIDLSLYTGSIMTYPIRREWEYE 240
QY      241 VIIVVEINGODLKMDCKEYNDKSIIVDSGTTNLRPKVFEAAVKSIIKAASSTKEKPPDG 300
DB      241 VIIVVEINGODLKMDCKEYNDKSIIVDSGTTNLRPKVFEAAVKSIIKAASSTKEKPPDG 300
QY      301 FWLGEOLVCMQAGTTPMNIIPVISLYLMGEVTNOSFRITLIPQOYLRPVEDVATSODDCY 360
DB      301 FWLGEOLVCMQAGTTPMNIIPVISLYLMGEVTNOSFRITLIPQOYLRPVEDVATSODDCY 360
QY      361 KEAISOSSSTGYMGAIVMEGFYVFPDRARKRIGFAVSAACHVHDEFRTAAVEGPFVTLDME 420
DB      361 KEAISOSSSTGYMGAIVMEGFYVFPDRARKRIGFAVSAACHVHDEFRTAAVEGPFVTLDME 420
QY      421 DCGYNIPTQDEST 433
```

```
DB      421 DCGYNIPTQDEST 433

RESULT 7
US-11-090-866-66
/ Sequence 66, Application US/11090866
/ Publication No. US20050164294A1
/ GENERAL INFORMATION:
/ APPLICANT: Anderson, John P.
/ APPLICANT: Basi, Gurigbal
/ APPLICANT: Doane, Minh Tam
/ APPLICANT: Frigon, Normand
/ APPLICANT: John, Varghese
/ APPLICANT: Power, Michael
/ APPLICANT: Sinha, Sukanto
/ APPLICANT: Tatsuno, Gwen
/ APPLICANT: Tung, Jay
/ APPLICANT: Wang, Shuwen
/ APPLICANT: McConlogue, Lisa
/ TITLE OF INVENTION: Beta-Secretase Enzyme Compositions and
/ FILE REFERENCE: 228-US-NEW2C12
/ CURRENT FILING DATE: 2005-03-25
/ PRIOR APPLICATION NUMBER: US/11/090,866
/ PRIOR FILING DATE: 2000-11-28
/ PRIOR APPLICATION NUMBER: US 09/723,722
/ PRIOR FILING DATE: 2000-02-10
/ PRIOR APPLICATION NUMBER: US 09/501,708
/ PRIOR FILING DATE: 2000-02-10
/ PRIOR APPLICATION NUMBER: US 09/471,669
/ PRIOR FILING DATE: 1999-12-24
/ PRIOR APPLICATION NUMBER: 60/119,571
/ PRIOR FILING DATE: 1999-02-10
/ PRIOR APPLICATION NUMBER: 60/139,172
/ PRIOR FILING DATE: 1999-06-15
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 66
/ LENGTH: 480
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-11-090-866-66

Query Match      99.1%; Score 2267; DB 20; Length 480;
Best Local Similarity 99.1%; Pred. No. 5.5e-209;
Matches 429; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      1 TOHGIRLPRLSGGAPLGLIEINLETDEPEEPGRGSGFVEMVDNLRGSGGGYVEMTV 60
DB      1 TOHGIRLPRLSGGAPLGLRLPRETDEPEEPGRGSGFVEMVDNLRGSGGGYVEMTV 60
QY      61 GSPPTLNLIVDTGSSNFVGAAPHPFLHRYYQROLSSTYRDLRGVYVPYTOGMEBEL 120
DB      61 GSPPTLNLIVDTGSSNFVGAAPHPFLHRYYQROLSSTYRDLRGVYVPYTOGMEBEL 120
QY      121 GTDLVSIPIHGPVNTYRANIAITTESDKPFINSNMEGLIGLAVAIAPDDSLPEFPSTL 180
DB      121 GTDLVSIPIHGPVNTYRANIAITTESDKPFINSNMEGLIGLAVAIAPDDSLPEFPSTL 180
QY      181 VQOTHPNLFSLQCGAGFPLNQSEVLASVGSMTIGSIDLSLYTGSIMTYPIRREWEYE 240
DB      181 VQOTHPNLFSLQCGAGFPLNQSEVLASVGSMTIGSIDLSLYTGSIMTYPIRREWEYE 240
QY      241 VIIVVEINGODLKMDCKEYNDKSIIVDSGTTNLRPKVFEAAVKSIIKAASSTKEKPPDG 300
DB      241 VIIVVEINGODLKMDCKEYNDKSIIVDSGTTNLRPKVFEAAVKSIIKAASSTKEKPPDG 300
QY      301 FWLGEOLVCMQAGTTPMNIIPVISLYLMGEVTNOSFRITLIPQOYLRPVEDVATSODDCY 360
DB      301 FWLGEOLVCMQAGTTPMNIIPVISLYLMGEVTNOSFRITLIPQOYLRPVEDVATSODDCY 360
QY      361 KEAISOSSSTGYMGAIVMEGFYVFPDRARKRIGFAVSAACHVHDEFRTAAVEGPFVTLDME 420
DB      361 KEAISOSSSTGYMGAIVMEGFYVFPDRARKRIGFAVSAACHVHDEFRTAAVEGPFVTLDME 420
```

Db 361 KFAISQSTGTGMGVINEGFYVFDPRARKIGFAVSACHVDEFRTAAVEGPFVTLDM 420  
Qy 421 DCGYINIPQTEST 433  
Db 421 DCGYINIPQTEST 433

RESULT 8  
US-11-069-377-66  
; Sequence 66, Application US/11069377  
; Publication No. US20050170489A1  
; GENERAL INFORMATION:  
; APPLICANT: Anderson, John P.  
; APPLICANT: Basal, Gurigbal  
; APPLICANT: Doane, Minh Tam  
; APPLICANT: Frigon, Normand  
; APPLICANT: John, Varghese  
; APPLICANT: Power, Michael  
; APPLICANT: Sinha, Sukanto  
; APPLICANT: Tatsuno, Gwen  
; APPLICANT: Tung, Jay  
; APPLICANT: Wang, Shuwen  
; APPLICANT: McConlogue, Lisa  
; TITLE OF INVENTION: Beta-Secretase Enzyme Compositions and  
; TITLE OF INVENTION: Methods  
; FILE REFERENCE: 228-US-NEW2C8  
; CURRENT APPLICATION NUMBER: US/11/069,377  
; CURRENT FILING DATE: 2005-02-28  
; PRIOR APPLICATION NUMBER: US 09/723,722  
; PRIOR FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: US 09/501,708  
; PRIOR FILING DATE: 2000-02-10  
; PRIOR APPLICATION NUMBER: 60/119,571  
; PRIOR FILING DATE: 1999-02-10  
; PRIOR APPLICATION NUMBER: 60/139,172  
; PRIOR FILING DATE: 1999-06-15  
; NUMBER OF SEQ ID NOS: 104  
; SOFTWARE: FaastSeq for Windows Version 4.0  
; SEQ ID NO 66  
; LENGTH: 480  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-069-377-66

Query Match 99.1%; Score 2267; DB 20; Length 480;  
Best Local Similarity 99.1%; Pred. No. 5.5e-209;  
Matches 429; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
Qy 1 TQHGIRLPRLSGAGAPLGLINLETDEBEPGRGSEFVEMVDNLRGSGGQYVEMTV 60  
1 TQHGIRLPRLSGAGAPLGLRLPRETDEBEPGRGSEFVEMVDNLRGSGGQYVEMTV 60  
Db 61 GSPPTLNTLIVDTGSSNFVGAAPHPFLHRYQROLSSTYRDLRKGVVYVYTGKMBGL 120  
61 GSPPTLNTLIVDTGSSNFVGAAPHPFLHRYQROLSSTYRDLRKGVVYVYTGKMBGL 120  
Qy 121 GTDLVSIPIHGPNTVYRANIAAITESDKFINGSNMGIIGLAVAEIARPDSDLEPFDSL 180  
121 GTDLVSIPIHGPNTVYRANIAAITESDKFINGSNMGIIGLAVAEIARPDSDLEPFDSL 180  
Db 121 GTDLVSIPIHGPNTVYRANIAAITESDKFINGSNMGIIGLAVAEIARPDSDLEPFDSL 180  
121 GTDLVSIPIHGPNTVYRANIAAITESDKFINGSNMGIIGLAVAEIARPDSDLEPFDSL 180  
Qy 181 VKQTHVNLFSIQLCGAGFPLNQSEVLASVGSMTIGIDHSLYTGSLMYTPIRRRMYE 240  
181 VKQTHVNLFSIQLCGAGFPLNQSEVLASVGSMTIGIDHSLYTGSLMYTPIRRRMYE 240  
Db 181 VKQTHVNLFSIQLCGAGFPLNQSEVLASVGSMTIGIDHSLYTGSLMYTPIRRRMYE 240  
181 VKQTHVNLFSIQLCGAGFPLNQSEVLASVGSMTIGIDHSLYTGSLMYTPIRRRMYE 240  
Qy 241 VTIIVVEINQDLKMDCKEYNDKSIYDSGTNLRLPKVFEAAVYSIKAASTKFPDG 300  
241 VTIIVVEINQDLKMDCKEYNDKSIYDSGTNLRLPKVFEAAVYSIKAASTKFPDG 300  
Db 241 VTIIVVEINQDLKMDCKEYNDKSIYDSGTNLRLPKVFEAAVYSIKAASTKFPDG 300  
241 VTIIVVEINQDLKMDCKEYNDKSIYDSGTNLRLPKVFEAAVYSIKAASTKFPDG 300  
Qy 301 FMIQEOLVCMQAGTTPWNIFFVYISLYLMGEVTNOSFRITILPQOYLRPVEDVATSODDCY 360  
301 FMIQEOLVCMQAGTTPWNIFFVYISLYLMGEVTNOSFRITILPQOYLRPVEDVATSODDCY 360  
Db 301 FMIQEOLVCMQAGTTPWNIFFVYISLYLMGEVTNOSFRITILPQOYLRPVEDVATSODDCY 360  
301 FMIQEOLVCMQAGTTPWNIFFVYISLYLMGEVTNOSFRITILPQOYLRPVEDVATSODDCY 360  
Qy 361 KFAISQSTGTGMGVINEGFYVFDPRARKIGFAVSACHVDEFRTAAVEGPFVTLDM 420

Db 361 KFAISQSTGTGMGVINEGFYVFDPRARKIGFAVSACHVDEFRTAAVEGPFVTLDM 420  
Qy 421 DCGYINIPQTEST 433  
Db 421 DCGYINIPQTEST 433

RESULT 9  
US-11-090-872-66  
; Sequence 66, Application US/11090872  
; Publication No. US20050177888A1  
; GENERAL INFORMATION:  
; APPLICANT: Anderson, John P.  
; APPLICANT: Basal, Gurigbal  
; APPLICANT: Doane, Minh Tam  
; APPLICANT: Frigon, Normand  
; APPLICANT: John, Varghese  
; APPLICANT: Power, Michael  
; APPLICANT: Sinha, Sukanto  
; APPLICANT: Tatsuno, Gwen  
; APPLICANT: Tung, Jay  
; APPLICANT: Wang, Shuwen  
; APPLICANT: McConlogue, Lisa  
; TITLE OF INVENTION: Beta-Secretase Enzyme Compositions and  
; TITLE OF INVENTION: Methods  
; FILE REFERENCE: 228-US-NEW2C11  
; CURRENT APPLICATION NUMBER: US/11/090,872  
; CURRENT FILING DATE: 2005-03-25  
; PRIOR APPLICATION NUMBER: US 09/723,722  
; PRIOR FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: US 09/501,708  
; PRIOR FILING DATE: 2000-02-10  
; PRIOR APPLICATION NUMBER: US 09/471,669  
; PRIOR FILING DATE: 1999-12-24  
; PRIOR APPLICATION NUMBER: 60/119,571  
; PRIOR FILING DATE: 1999-02-10  
; PRIOR APPLICATION NUMBER: 60/139,172  
; PRIOR FILING DATE: 1999-06-15  
; NUMBER OF SEQ ID NOS: 104  
; SOFTWARE: FaastSeq for Windows Version 4.0  
; SEQ ID NO 66  
; LENGTH: 480  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-090-872-66

Query Match 99.1%; Score 2267; DB 20; Length 480;  
Best Local Similarity 99.1%; Pred. No. 5.5e-209;  
Matches 429; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
Qy 1 TQHGIRLPRLSGAGAPLGLINLETDEBEPGRGSEFVEMVDNLRGSGGQYVEMTV 60  
1 TQHGIRLPRLSGAGAPLGLRLPRETDEBEPGRGSEFVEMVDNLRGSGGQYVEMTV 60  
Db 61 GSPPTLNTLIVDTGSSNFVGAAPHPFLHRYQROLSSTYRDLRKGVVYVYTGKMBGL 120  
61 GSPPTLNTLIVDTGSSNFVGAAPHPFLHRYQROLSSTYRDLRKGVVYVYTGKMBGL 120  
Qy 121 GTDLVSIPIHGPNTVYRANIAAITESDKFINGSNMGIIGLAVAEIARPDSDLEPFDSL 180  
121 GTDLVSIPIHGPNTVYRANIAAITESDKFINGSNMGIIGLAVAEIARPDSDLEPFDSL 180  
Db 121 GTDLVSIPIHGPNTVYRANIAAITESDKFINGSNMGIIGLAVAEIARPDSDLEPFDSL 180  
121 GTDLVSIPIHGPNTVYRANIAAITESDKFINGSNMGIIGLAVAEIARPDSDLEPFDSL 180  
Qy 181 VKQTHVNLFSIQLCGAGFPLNQSEVLASVGSMTIGIDHSLYTGSLMYTPIRRRMYE 240  
181 VKQTHVNLFSIQLCGAGFPLNQSEVLASVGSMTIGIDHSLYTGSLMYTPIRRRMYE 240  
Db 181 VKQTHVNLFSIQLCGAGFPLNQSEVLASVGSMTIGIDHSLYTGSLMYTPIRRRMYE 240  
181 VKQTHVNLFSIQLCGAGFPLNQSEVLASVGSMTIGIDHSLYTGSLMYTPIRRRMYE 240  
Qy 241 VTIIVVEINQDLKMDCKEYNDKSIYDSGTNLRLPKVFEAAVYSIKAASTKFPDG 300  
241 VTIIVVEINQDLKMDCKEYNDKSIYDSGTNLRLPKVFEAAVYSIKAASTKFPDG 300  
Db 241 VTIIVVEINQDLKMDCKEYNDKSIYDSGTNLRLPKVFEAAVYSIKAASTKFPDG 300  
241 VTIIVVEINQDLKMDCKEYNDKSIYDSGTNLRLPKVFEAAVYSIKAASTKFPDG 300  
Qy 301 FMIQEOLVCMQAGTTPWNIFFVYISLYLMGEVTNOSFRITILPQOYLRPVEDVATSODDCY 360  
301 FMIQEOLVCMQAGTTPWNIFFVYISLYLMGEVTNOSFRITILPQOYLRPVEDVATSODDCY 360  
Db 301 FMIQEOLVCMQAGTTPWNIFFVYISLYLMGEVTNOSFRITILPQOYLRPVEDVATSODDCY 360  
301 FMIQEOLVCMQAGTTPWNIFFVYISLYLMGEVTNOSFRITILPQOYLRPVEDVATSODDCY 360

Db 301 FWLGEOLVCMQAGTTPMNI FVVISLYLMEVNTQSFRTITL PQQYLRPVEDVATSDDCY 360  
QY 361 KFAISQSTGTVMGAVIMEGFYVFPDRARRKIGFAVSAACHVDEFRITAAVEGPFVTL DME 420  
Db 361 KFAISQSTGTVMGAVIMEGFYVFPDRARRKIGFAVSAACHVDEFRITAAVEGPFVTL DME 420  
QY 421 DCGYNIPQTDTEST 433  
Db 421 DCGYNIPQTDTEST 433

RESULT 10  
US-09-796-264-2  
Sequence 2, Application US/09796264  
Patent No. US20020049303A1  
GENERAL INFORMATION:  
APPLICANT: Tang, Jordan J.N.  
APPLICANT: Lin, Xinli  
TITLE OF INVENTION: Catalytically Active Recombinant Memapsin and Methods  
FILE REFERENCE: OMRP 179  
CURRENT APPLICATION NUMBER: US/09/796,264  
PRIOR FILING DATE: 2001-02-28  
PRIOR APPLICATION NUMBER: 09/604,608  
PRIOR FILING DATE: 2000-06-27  
PRIOR APPLICATION NUMBER: 60/168,060  
PRIOR FILING DATE: 1999-11-30  
PRIOR APPLICATION NUMBER: 60/177,836  
PRIOR FILING DATE: 2000-01-25  
PRIOR APPLICATION NUMBER: 60/178,368  
PRIOR FILING DATE: 2000-01-27  
PRIOR APPLICATION NUMBER: 60/210,292  
PRIOR FILING DATE: 2000-06-08  
NUMBER OF SEQ ID NOS: 31  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 488  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: Purified Memapsin 2  
OTHER INFORMATION: Amino Acids 28-48 are remnant putative propeptide  
OTHER INFORMATION: Residues  
OTHER INFORMATION: Amino Acids 58-61, 78, 80, 82-83, 116, 118-121,  
OTHER INFORMATION: 156, 166, 174, 246, 274, 276, 278-281, 283, and  
OTHER INFORMATION: 376-377 are residues in contact with the OMR9-2  
OTHER INFORMATION: Inhibitor  
OTHER INFORMATION: Amino acids 54-57, 61-68, 73-80, 86-89, 109-111,  
OTHER INFORMATION: 113-118, 123-134, 143-154, 165-168, 198-202, and  
OTHER INFORMATION: 220-224 are N-lobe Beta Strands  
OTHER INFORMATION: Amino Acids 184-191 and 210-217 are N-lobe Helices  
OTHER INFORMATION: Amino acids 237-240, 247-249, 251-256, 259-260,  
OTHER INFORMATION: 273-275, 282-285, 316-318, 331-336, 342-348,  
OTHER INFORMATION: 354-357, 366-370, 372-375, 380-383, 390-395,  
OTHER INFORMATION: 400-405, and 418-420 are C-lobe Beta Strands  
OTHER INFORMATION: Amino Acids 286-299, 307-310, 350-353, 384-387,  
OTHER INFORMATION: and 427-431 are C-lobe Helices  
US-09-796-264-2

Query Match 99.1%; Score 2267; DB 9; Length 488;  
Best Local Similarity 99.1%; Pred No. 5,6e-209;  
Matches 429; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 TQHGRLRLRSGLGAPLGLLEINLTDEPEBPGRGSFVEMVNLNGSKSGQGYVEMTV 60  
Db 9 TQHGRLRLRSGLGAPLGLRLPRETDEPEBPGRGSFVEMVNLNGSKSGQGYVEMTV 68  
QY 61 GSPPTLNIIVDTGSSNFAVGAAPRPHRRYYQQLSTTYDLDKGYVYPYTGKWEGL 120  
Db 69 GSPPTLNIIVDTGSSNFAVGAAPRPHRRYYQQLSTTYDLDKGYVYPYTGKWEGL 128  
QY 121 GTDLVSIPIHGRNVYTRANIAAITESDKFFINGSNMEGLGLAVALIARPDOSLEPPDSL 180

Db 129 GTDLVSIPIHGRNVYTRANIAAITESDKFFINGSNMEGLGLAVALIARPDOSLEPPDSL 188  
QY 181 VQOTHPNLFSIQLGAGPPLNQSVLASVGSMTIGSIDSLYTGSLMYTPIRREWYE 240  
Db 189 VQOTHPNLFSIQLGAGPPLNQSVLASVGSMTIGSIDSLYTGSLMYTPIRREWYE 248  
QY 241 VTIIRVEINGDPLKMDCKEYVYDKSIYDSGTNNLRPKVFEAAVKSISKASSTEKPPDG 300  
Db 249 VTIIRVEINGDPLKMDCKEYVYDKSIYDSGTNNLRPKVFEAAVKSISKASSTEKPPDG 308  
QY 301 FWLGEOLVCMQAGTTPMNI FVVISLYLMEVNTQSFRTITL PQQYLRPVEDVATSDDCY 360  
Db 309 FWLGEOLVCMQAGTTPMNI FVVISLYLMEVNTQSFRTITL PQQYLRPVEDVATSDDCY 368  
QY 361 KFAISQSTGTVMGAVIMEGFYVFPDRARRKIGFAVSAACHVDEFRITAAVEGPFVTL DME 420  
Db 369 KFAISQSTGTVMGAVIMEGFYVFPDRARRKIGFAVSAACHVDEFRITAAVEGPFVTL DME 428  
QY 421 DCGYNIPQTDTEST 433  
Db 429 DCGYNIPQTDTEST 441

RESULT 11  
US-09-845-226-2  
Sequence 2, Application US/09845226  
Patent No. US20020115600A1  
GENERAL INFORMATION:  
APPLICANT: Tang, Jordan J.N.  
APPLICANT: Hong, Lin  
TITLE OF INVENTION: Inhibitors of Memapsin 2 and Use Thereof  
FILE REFERENCE: OMRP 182  
CURRENT APPLICATION NUMBER: US/09/845,226  
PRIOR FILING DATE: 2001-04-30  
PRIOR APPLICATION NUMBER: 09/603,713  
PRIOR FILING DATE: 2000-06-27  
PRIOR APPLICATION NUMBER: 60/168,060  
PRIOR FILING DATE: 1999-11-30  
PRIOR APPLICATION NUMBER: 60/177,836  
PRIOR FILING DATE: 2000-01-25  
PRIOR APPLICATION NUMBER: 60/178,368  
PRIOR FILING DATE: 2000-01-27  
PRIOR APPLICATION NUMBER: 60/210,292  
PRIOR FILING DATE: 2000-06-08  
NUMBER OF SEQ ID NOS: 31  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 488  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: Purified Memapsin 2  
OTHER INFORMATION: Amino Acids 28-48 are remnant putative propeptide  
OTHER INFORMATION: Residues  
OTHER INFORMATION: Amino Acids 58-61, 78, 80, 82-83, 116, 118-121,  
OTHER INFORMATION: 156, 166, 174, 246, 274, 276, 278-281, 283, and  
OTHER INFORMATION: 376-377 are residues in contact with the OMR9-2  
OTHER INFORMATION: Inhibitor  
OTHER INFORMATION: Amino acids 54-57, 61-68, 73-80, 86-89, 109-111,  
OTHER INFORMATION: 113-118, 123-134, 143-154, 165-168, 198-202, and  
OTHER INFORMATION: 220-224 are N-lobe Beta Strands  
OTHER INFORMATION: Amino Acids 184-191 and 210-217 are N-lobe Helices  
OTHER INFORMATION: Amino acids 237-240, 247-249, 251-256, 259-260,  
OTHER INFORMATION: 273-275, 282-285, 316-318, 331-336, 342-348,  
OTHER INFORMATION: 354-357, 366-370, 372-375, 380-383, 390-395,  
OTHER INFORMATION: 400-405, and 418-420 are C-lobe Beta Strands  
OTHER INFORMATION: Amino Acids 286-299, 307-310, 350-353, 384-387,  
OTHER INFORMATION: and 427-431 are C-lobe Helices  
US-09-845-226-2

Query Match 99.1%; Score 2267; DB 9; Length 488;



```

Best Local Similarity 99.1%; Pred. No. 5.6e-209;
Matches 429; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 TQHGIRLPLRSGLGAPLGLINLETDEEPEEPRGRSGFVEMVDNLRGSGGYVEMTV 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 9 TQHGIRLPLRSGLGAPLGLINLETDEEPEEPRGRSGFVEMVDNLRGSGGYVEMTV 68

QY 61 GSPPTLNLVDTGSSNFAVGAAPHPFLHRYQRLSSTYRDLRGVVPYTOGKMEGL 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 69 GSPPTLNLVDTGSSNFAVGAAPHPFLHRYQRLSSTYRDLRGVVPYTOGKMEGL 128

QY 121 GTDIVSIPHPNVTVRANIAATTESDKFFINGSMNEGILGLAYAEIARPDLSLEPPDSL 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 129 GTDIVSIPHPNVTVRANIAATTESDKFFINGSMNEGILGLAYAEIARPDLSLEPPDSL 188

QY 181 VKQTHVPMFLSLOLCAGAPFLNQLSEVLASVGSMTIGIDHSLYTGSLWYTPIRREWE 240
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 189 VKQTHVPMFLSLOLCAGAPFLNQLSEVLASVGSMTIGIDHSLYTGSLWYTPIRREWE 248

QY 241 VIIIRVEINGODLKMDCKEYNVDKSIYDSGTTNLRPKVFEAAVKSIAAASSTKFPDG 300
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 249 VIIIRVEINGODLKMDCKEYNVDKSIYDSGTTNLRPKVFEAAVKSIAAASSTKFPDG 308

QY 301 FWMGEOLVCMQAGTTPMNIFFPVISLYLMEGVNOSFRITLLPOQYLRPEVDVATSODDCY 360
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 309 FWMGEOLVCMQAGTTPMNIFFPVISLYLMEGVNOSFRITLLPOQYLRPEVDVATSODDCY 368

QY 361 KPAISQSTGTVMGAVIMGFYVDFRARKRIGFAVSACHVDEFTAAVEGPFVTLDM 420
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 369 KPAISQSTGTVMGAVIMGFYVDFRARKRIGFAVSACHVDEFTAAVEGPFVTLDM 428

QY 421 DCGYNIPQTDST 433
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 429 DCGYNIPQTDST 441

RESULT 12
US-09-795-903A-2
; Sequence 2, Application US/09795903A
; Patent No. US20020164760A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Jordan J.N.
; APPLICANT: Lin, Xindi
; APPLICANT: Koelsch, Gerald
; TITLE OF INVENTION: Catalytically Active Recombinant Memapsin and Methods
; FILE REFERENCE: OMRP 179
; CURRENT APPLICATION NUMBER: US/09/795,903A
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 09/604,608
; PRIOR FILING DATE: 2000-06-27
; PRIOR APPLICATION NUMBER: 60/168,060
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: 60/177,836
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 60/178,368
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: 60/210,292
; PRIOR FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 488
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Purified Memapsin 2
; OTHER INFORMATION: Amino Acids 28-48 are remnant putative propeptide
; OTHER INFORMATION: residues
; OTHER INFORMATION: Amino Acids 58-61, 78, 80, 82-83, 116, 118-121,
; OTHER INFORMATION: 156, 166, 174, 246, 274, 276, 278-281, 283, and
; OTHER INFORMATION: 376-377 are residues in contact with the OMRP-2
; OTHER INFORMATION: inhibitor

```

```

OTHER INFORMATION: Amino acids 54-57, 61-68, 73-80, 86-89, 109-111,
; OTHER INFORMATION: 113-118, 123-134, 143-154, 165-168, 198-202, and
; OTHER INFORMATION: 220-224 are N-lobe Beta Strands
; OTHER INFORMATION: 184-191 and 210-217 are N-lobe Helices
; OTHER INFORMATION: Amino acids 237-240, 247-249, 251-256, 259-260,
; OTHER INFORMATION: 273-275, 282-285, 316-318, 331-336, 342-348,
; OTHER INFORMATION: 354-357, 366-370, 372-375, 380-383, 390-395,
; OTHER INFORMATION: 400-405, and 418-420 are C-lobe Beta Strands
; OTHER INFORMATION: Amino Acids 286-299, 307-310, 350-353, 384-387,
; OTHER INFORMATION: and 427-431 are C-lobe Helices
US-09-795-903A-2

Query Match 99.1%; Score 2267; DB 9; Length 488;
Best Local Similarity 99.1%; Pred. No. 5.6e-209;
Matches 429; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 TQHGIRLPLRSGLGAPLGLINLETDEEPEEPRGRSGFVEMVDNLRGSGGYVEMTV 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 9 TQHGIRLPLRSGLGAPLGLINLETDEEPEEPRGRSGFVEMVDNLRGSGGYVEMTV 68

QY 61 GSPPTLNLVDTGSSNFAVGAAPHPFLHRYQRLSSTYRDLRGVVPYTOGKMEGL 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 69 GSPPTLNLVDTGSSNFAVGAAPHPFLHRYQRLSSTYRDLRGVVPYTOGKMEGL 128

QY 121 GTDIVSIPHPNVTVRANIAATTESDKFFINGSMNEGILGLAYAEIARPDLSLEPPDSL 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 129 GTDIVSIPHPNVTVRANIAATTESDKFFINGSMNEGILGLAYAEIARPDLSLEPPDSL 188

QY 181 VKQTHVPMFLSLOLCAGAPFLNQLSEVLASVGSMTIGIDHSLYTGSLWYTPIRREWE 240
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 189 VKQTHVPMFLSLOLCAGAPFLNQLSEVLASVGSMTIGIDHSLYTGSLWYTPIRREWE 248

QY 241 VIIIRVEINGODLKMDCKEYNVDKSIYDSGTTNLRPKVFEAAVKSIAAASSTKFPDG 300
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 249 VIIIRVEINGODLKMDCKEYNVDKSIYDSGTTNLRPKVFEAAVKSIAAASSTKFPDG 308

QY 301 FWMGEOLVCMQAGTTPMNIFFPVISLYLMEGVNOSFRITLLPOQYLRPEVDVATSODDCY 360
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 309 FWMGEOLVCMQAGTTPMNIFFPVISLYLMEGVNOSFRITLLPOQYLRPEVDVATSODDCY 368

QY 361 KPAISQSTGTVMGAVIMGFYVDFRARKRIGFAVSACHVDEFTAAVEGPFVTLDM 420
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 369 KPAISQSTGTVMGAVIMGFYVDFRARKRIGFAVSACHVDEFTAAVEGPFVTLDM 428

QY 421 DCGYNIPQTDST 433
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 429 DCGYNIPQTDST 441

RESULT 13
US-10-032-818-2
; Sequence 2, Application US/10032818
; Publication No. US20030092629A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Jordan J.N.
; APPLICANT: Koelsch, Gerald
; APPLICANT: Ghosh, Arun K.
; TITLE OF INVENTION: Inhibitors of Memapsin 2 and Use Thereof
; FILE REFERENCE: 2932.1006-007
; CURRENT APPLICATION NUMBER: US/10/032,818
; CURRENT FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: US 60/275,756
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: US 60/258,705
; PRIOR FILING DATE: 2000-12-28
; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 488
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-032-818-2

```

Query Match 99.1%; Score 2267; DB 14; Length 488;  
Best Local Similarity 99.1%; Pred. No. 5,66-209;  
Matches 429; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 TQHGIRLPRLSGGLGAPLGLINLETDEEPEERGRSGFVEMVDNLGKSGGGYVEMTV 60  
DB 9 TQHGIRLPRLSGGLGAPLGLRLPRETDEEPEERGRSGFVEMVDNLGKSGGGYVEMTV 68  
QY 61 GSPQTNIILVDTGSSNFAVGAAPHPLHRYRQRLSSTYRDLRKGYVYPTQKMEGEL 120  
DB 69 GSPQTNIILVDTGSSNFAVGAAPHPLHRYRQRLSSTYRDLRKGYVYPTQKMEGEL 128  
QY 121 GTDLVSIPIHGNVTYVRANIAITSDKPFINGSNWEGILGLAYAEIARPPDSLEPPDSL 180  
DB 129 GTDLVSIPIHGNVTYVRANIAITSDKPFINGSNWEGILGLAYAEIARPPDSLEPPDSL 188  
QY 181 VKQTHVNLFSLOLCGAGFPLNOSVLA SVGSMITIGIDHSLYTSLWTPPIRRMEY 240  
DB 189 VKQTHVNLFSLOLCGAGFPLNOSVLA SVGSMITIGIDHSLYTSLWTPPIRRMEY 248  
QY 241 VIIVRVINGQDLKMDCKEYNDKSIYDSGTTNLRPLPKVFEAAVKSIAASSTKPPDG 300  
DB 249 VIIVRVINGQDLKMDCKEYNDKSIYDSGTTNLRPLPKVFEAAVKSIAASSTKPPDG 308  
QY 301 FMIGEOLVCMQAGTTPNIPFVLSLYMGEVNTQSFRTITLPOQYLRPEVAVTSODDCY 360  
DB 309 FMIGEOLVCMQAGTTPNIPFVLSLYMGEVNTQSFRTITLPOQYLRPEVAVTSODDCY 368  
QY 361 KFAISGSGTGVGAVIMESGFYVFPDRAKRIGFAVSAHVDEPRTAAVEGPFVTLDM 420  
DB 369 KFAISGSGTGVGAVIMESGFYVFPDRAKRIGFAVSAHVDEPRTAAVEGPFVTLDM 428  
QY 421 DCGYNIPQTDST 433  
DB 429 DCGYNIPQTDST 441

RESULT 14  
US-10-820-953-2  
Sequence 2, Application US/10820953  
Publication No. US20040167075A1  
GENERAL INFORMATION:  
APPLICANT: Tang, Jordan J.N.  
APPLICANT: Hong, Lin  
APPLICANT: Ghosh, Arun K.  
TITLE OF INVENTION: Inhibitors of Memapsin 2 and Use Thereof  
FILE REFERENCE: OMRP 182  
CURRENT APPLICATION NUMBER: US/10/820,953  
PRIOR FILING DATE: 2004-04-08  
PRIOR APPLICATION NUMBER: US/09/603,713  
PRIOR FILING DATE: 2000-06-27  
PRIOR APPLICATION NUMBER: 60/141,363  
PRIOR FILING DATE: 1999-06-28  
PRIOR APPLICATION NUMBER: 60/168,060  
PRIOR FILING DATE: 1999-11-30  
PRIOR APPLICATION NUMBER: 60/177,836  
PRIOR FILING DATE: 2000-01-25  
PRIOR APPLICATION NUMBER: 60/178,368  
PRIOR FILING DATE: 2000-01-27  
PRIOR APPLICATION NUMBER: 60/210,292  
PRIOR FILING DATE: 2000-06-06  
NUMBER OF SEQ ID NOS: 31  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 488  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: Purified Memapsin 2  
OTHER INFORMATION: Amino Acids 28-48 are remnant putative propeptide  
OTHER INFORMATION: Residues  
OTHER INFORMATION: Amino Acids 58-61, 78, 80, 82-83, 116, 118-121,  
OTHER INFORMATION: 156, 166, 174, 246, 274, 276, 278-281, 283, and

OTHER INFORMATION: 376-377 are residues in contact with the OM99-2  
OTHER INFORMATION: Inhibitor  
OTHER INFORMATION: Amino acids 54-57, 61-68, 73-80, 86-89, 109-111,  
OTHER INFORMATION: 113-118, 123-134, 143-154, 165-168, 198-202, and  
OTHER INFORMATION: 220-224 are N-lobe Beta Strands  
OTHER INFORMATION: Amino acids 184-191 and 210-217 are N-lobe Helices  
OTHER INFORMATION: Amino acids 237-240, 247-249, 251-256, 259-260,  
OTHER INFORMATION: 273-275, 282-285, 316-318, 331-336, 342-348,  
OTHER INFORMATION: 354-357, 366-370, 372-375, 380-383, 390-395,  
OTHER INFORMATION: 400-405, and 418-420 are C-lobe Beta Strands  
OTHER INFORMATION: Amino Acids 286-299, 307-310, 350-353, 364-387,  
OTHER INFORMATION: and 427-431 are C-lobe Helices  
US-10-820-953-2

Query Match 99.1%; Score 2267; DB 16; Length 488;  
Best Local Similarity 99.1%; Pred. No. 5,66-209;  
Matches 429; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 TQHGIRLPRLSGGLGAPLGLINLETDEEPEERGRSGFVEMVDNLGKSGGGYVEMTV 60  
DB 9 TQHGIRLPRLSGGLGAPLGLRLPRETDEEPEERGRSGFVEMVDNLGKSGGGYVEMTV 68  
QY 61 GSPQTNIILVDTGSSNFAVGAAPHPLHRYRQRLSSTYRDLRKGYVYPTQKMEGEL 120  
DB 69 GSPQTNIILVDTGSSNFAVGAAPHPLHRYRQRLSSTYRDLRKGYVYPTQKMEGEL 128  
QY 121 GTDLVSIPIHGNVTYVRANIAITSDKPFINGSNWEGILGLAYAEIARPPDSLEPPDSL 180  
DB 129 GTDLVSIPIHGNVTYVRANIAITSDKPFINGSNWEGILGLAYAEIARPPDSLEPPDSL 188  
QY 181 VKQTHVNLFSLOLCGAGFPLNOSVLA SVGSMITIGIDHSLYTSLWTPPIRRMEY 240  
DB 189 VKQTHVNLFSLOLCGAGFPLNOSVLA SVGSMITIGIDHSLYTSLWTPPIRRMEY 248  
QY 241 VIIVRVINGQDLKMDCKEYNDKSIYDSGTTNLRPLPKVFEAAVKSIAASSTKPPDG 300  
DB 249 VIIVRVINGQDLKMDCKEYNDKSIYDSGTTNLRPLPKVFEAAVKSIAASSTKPPDG 308  
QY 301 FMIGEOLVCMQAGTTPNIPFVLSLYMGEVNTQSFRTITLPOQYLRPEVAVTSODDCY 360  
DB 309 FMIGEOLVCMQAGTTPNIPFVLSLYMGEVNTQSFRTITLPOQYLRPEVAVTSODDCY 368  
QY 361 KFAISGSGTGVGAVIMESGFYVFPDRAKRIGFAVSAHVDEPRTAAVEGPFVTLDM 420  
DB 369 KFAISGSGTGVGAVIMESGFYVFPDRAKRIGFAVSAHVDEPRTAAVEGPFVTLDM 428  
QY 421 DCGYNIPQTDST 433  
DB 429 DCGYNIPQTDST 441

RESULT 15  
US-10-773-754-2  
Sequence 2, Application US/10773754  
Publication No. US20040220079A1  
GENERAL INFORMATION:  
APPLICANT: Tang, Jordan J. N.  
APPLICANT: Koelsch, Gerald  
APPLICANT: Ghosh, Arun K.  
APPLICANT: Hong, Lin  
APPLICANT: The Board of Trustees of the University of Illinois  
TITLE OF INVENTION: Inhibitors of Memapsin 2 and Use Thereof  
FILE REFERENCE: 022266-000930US  
CURRENT APPLICATION NUMBER: US/10/773,754  
PRIOR FILING DATE: 2004-02-06  
PRIOR APPLICATION NUMBER: US 60/141,363  
PRIOR FILING DATE: 1999-06-28  
PRIOR APPLICATION NUMBER: US 60/168,060  
PRIOR FILING DATE: 1999-11-30  
PRIOR APPLICATION NUMBER: US 60/177,836  
PRIOR FILING DATE: 2000-01-25  
PRIOR APPLICATION NUMBER: US 60/178,368

/ PRIOR FILING DATE: 2000-01-27  
/ PRIOR APPLICATION NUMBER: US 60/210,292  
/ PRIOR FILING DATE: 2000-06-08  
/ PRIOR APPLICATION NUMBER: US 09/603,713  
/ PRIOR FILING DATE: 2000-06-27  
/ PRIOR APPLICATION NUMBER: US 09/845,226  
/ PRIOR FILING DATE: 2001-04-30  
/ NUMBER OF SEQ ID NOS: 39  
/ SOFTWARE: PatentIn Ver. 2.1  
/ SEQ ID NO 2  
/ LENGTH: 488  
/ TYPE: PRT  
/ ORGANISM: Homo sapiens  
/ FEATURE:  
/ OTHER INFORMATION: purified memapsin 2, aspartic proteainase 2 (ASP2)  
/ FEATURE:  
/ OTHER INFORMATION: amino acids 28-48 are remnant putative propeptide  
/ OTHER INFORMATION: residues  
/ FEATURE:  
/ OTHER INFORMATION: amino acids 54-57, 61-68, 73-80, 86-89, 109-111,  
/ OTHER INFORMATION: 113-118, 123-134, 143-154, 165-168, 198-202, and  
/ OTHER INFORMATION: 220-224 are N-lobe beta strands  
/ FEATURE:  
/ OTHER INFORMATION: amino acids 58-61, 78, 80, 82-83, 116, 118-121,  
/ OTHER INFORMATION: 156, 166, 174, 246, 274, 276, 278-281, 283, and  
/ OTHER INFORMATION: 376-377 are residues in contact with the OM99-2  
/ OTHER INFORMATION: inhibitor  
/ FEATURE:  
/ OTHER INFORMATION: amino acids 184-191 and 210-217 are N-lobe helices  
/ FEATURE:  
/ OTHER INFORMATION: amino acids 237-240, 247-249, 251-256, 259-260,  
/ OTHER INFORMATION: 273-275, 282-285, 316-318, 331-336, 342-348,  
/ OTHER INFORMATION: 354-357, 366-370, 372-375, 380-383, 390-395,  
/ OTHER INFORMATION: 400-405, and 418-420 are C-lobe beta strands  
/ FEATURE:  
/ OTHER INFORMATION: amino acids 286-299, 307-310, 350-353, 384-387,  
/ OTHER INFORMATION: and 427-431 are C-lobe helices  
/ US-10-773-754-2

Query Match 99.1%; Score 2267; DB 16; Length 488;  
Best Local Similarity 99.1%; Pred. No. 5,6e-209;  
Matches 429; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 TQHGIRLPPLRSGLGAPLGLINLETDPEPEPEGRGSPVEMVDNLKSGGQGYVEMTV 60  
DB 9 TQHGIRLPPLRSGLGAPLGLIRLPRETDEPEPEGRGSPVEMVDNLKSGGQGYVEMTV 68  
QY 61 GSPPTLNIIVDTGSSNFAVGAAPHPLHRYQRLSSTYRDLKGYVVPYTGKWEGL 120  
DB 69 GSPPTLNIIVDTGSSNFAVGAAPHPLHRYQRLSSTYRDLKGYVVPYTGKWEGL 128  
QY 121 GTPLVSLPHGPNTVAVANIAITESPDKFFINGSNWEGILGLAYAEIARPPDSLPPFDSL 180  
DB 129 GTPLVSLPHGPNTVAVANIAITESPDKFFINGSNWEGILGLAYAEIARPPDSLPPFDSL 188  
QY 181 VKQTHVNPILFSLQICGAFPLNOSVVLASVGGSMIIIGIDHSLDTGSLWYTPIRREWYE 240  
DB 189 VKQTHVNPILFSLQICGAFPLNOSVVLASVGGSMIIIGIDHSLDTGSLWYTPIRREWYE 248  
QY 241 VILVVEINQDILMDCKEYNDKSIYDSGTTNLRLPKVFPAVAVKSIKAASTKEKPPDG 300  
DB 249 VILVVEINQDILMDCKEYNDKSIYDSGTTNLRLPKVFPAVAVKSIKAASTKEKPPDG 308  
QY 301 FMIGBOLVCMQAGTTPNIFPVISLYLMGEVNTOSFRITILPOOYLRPVEDVATSODDCY 360  
DB 309 FMIGBOLVCMQAGTTPNIFPVISLYLMGEVNTOSFRITILPOOYLRPVEDVATSODDCY 368  
QY 361 KPAISGSSTGTWGAIVIMSGFYVVFDRARRKIGFPAVSACHVHDEFRTAAVEGPPVTLDM 420  
DB 369 KPAISGSSTGTWGAIVIMSGFYVVFDRARRKIGFPAVSACHVHDEFRTAAVEGPPVTLDM 428  
QY 421 DCGYNIPTDEST 433  
|||||

DB 429 DCGYNIPTDEST 441

Search completed: August 22, 2005, 00:54:04  
Job time: 1528 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 21, 2005, 06:58:03 ; Search time 125 Seconds  
(without alignment)  
1339.738 Million cell updates/sec

Title: US-10-726-967a-84

Perfect score: 2288  
Sequence: 1 TQHGIRLPLRSGLGAPPLGL.....FTYLMEDCGINIPQTDST 433

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

- 1: geneseqp19808:\*
- 2: geneseqp19908:\*
- 3: geneseqp20008:\*
- 4: geneseqp20018:\*
- 5: geneseqp20028:\*
- 6: geneseqp20038:\*
- 7: geneseqp20038:\*
- 8: geneseqp20048:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2288	100.0	433	8	ADP83957 Human BAC
2	2286	99.9	433	8	ADP83954 Human BAC
3	2273	99.3	433	8	ADP83951 Human BAC
4	2267	99.1	435	8	AD164641 Mature hu
5	2267	99.1	456	6	ABR61929 Human pro
6	2267	99.1	488	4	AAB61334 Memapsin
7	2267	99.1	488	4	AAB66572 Human mem
8	2267	99.1	488	5	ABG78372 Human pro
9	2267	99.1	488	5	AU99488 Human pro
10	2267	99.1	488	6	ABG76101 Human par
11	2267	99.1	488	6	AD74817 Human mem
12	2267	99.1	501	3	AD194767 Human bet
13	2267	99.1	501	3	AB07896 Antno aci
14	2267	99.1	501	3	ABG78374 Human pre
15	2267	99.1	501	6	ABR61928 Human mem
16	2267	99.1	501	7	ADP81560 Human BAC
17	2267	99.1	501	7	ADL18184 Human APP
18	2267	99.1	501	8	ADG86751 Human APP
19	2267	99.1	501	8	ADJ94316 Human asp
20	2267	99.1	501	8	AD050412 Human asp
21	2267	99.1	501	8	AD017468 Human sof
22	2267	99.1	501	8	ADP83876 Human BAC
23	2267	99.1	501	8	ADR75325 Human asp
24	2267	99.1	501	4	AAB61335 T7 promot
25	2267	99.1	503	4	AAB66573 Human pro

26	2267	99.1	503	5	ABG78373 Human pre
27	2267	99.1	503	6	AU99489 Pro-mem
28	2267	99.1	503	6	ABG76102 Human mem
29	2267	99.1	503	6	AD74818 Plasmid C
30	2267	99.1	509	5	AAM52697 FLAG-tag
31	2267	99.1	509	8	ADJ71858 Human pro
32	2267	99.1	509	8	ADP74535 Human ind
33	2267	98.9	453	3	AAV88438 Modified
34	2267	98.9	453	4	AAU07215 Human asp
35	2267	98.9	453	4	AAE10642 Human-asp
36	2267	98.9	453	4	AAE06872 Human-asp
37	2267	98.9	453	4	AAE02594 Human-asp
38	2267	98.9	453	4	AAU06616 Human-pro
39	2267	98.9	453	5	ABR78603 Human-asp
40	2267	98.9	453	8	ADJ94342 Human-asp
41	2267	98.9	453	8	AD050438 Human-asp
42	2267	98.9	453	8	ADR75351 Human-asp
43	2267	98.9	455	8	ADJ57773 BACE WT P
44	2267	98.9	459	3	AAV88439 Modified
45	2267	98.9	459	4	AAU07216 Human asp

## ALIGNMENTS

RESULT 1  
ID ADP83957 standard; protein; 433 AA.

XX AC ADP83957;

XX DT 23-SEP-2004 (first entry)

XX DE Human BACE1 mutant amino acid sequence SEQ ID NO:84.

XX KM human; beta-site amyloid precursor protein cleaving enzyme 1;

XX KW beta-site APP cleaving enzyme 1; BACE1; BACE1 isoform A; chromosome 11;

XX KW prodomain; engineered cleavage site; protease domain; neuroprotective;

XX KW nootropic; gene therapy; Alzheimer's disease; Down's syndrome; mutant.

XX OS Homo sapiens.

XX OS Synthetic.

XX PN MO2004056962-A2.

XX PD 08-JUL-2004.

XX PF 02-DEC-2003; 2003WO-US038314.

XX PR 04-DEC-2002; 2002US-0430984P.

XX PA (SUNE-) SUNESIS PHARM INC.

XX PI Ballinger M;

XX DR WPI; 2004-S07703/48.

XX PT New polypeptides for producing homogeneously processed preparations of

XX PT beta site amyloid precursor protein-cleaving enzyme comprises a

XX PT prodomain, an engineered cleavage site and a protease domain.

XX PS Example 1; SEQ ID NO 84; 40pp; English.

XX CC The present invention describes a polypeptide (1) comprising in order

XX CC from the N-terminus to the C-terminus: (a) a prodomain comprising at

XX CC least 6 contiguous amino acids of the 16 amino acid sequence of SEQ ID

XX CC NO:3 (ADP83877), comprising residues 22-37 of SEQ ID NO:1 (ADP83876) which

XX CC is the longest isoform of human beta-site amyloid precursor protein (APP)

XX CC cleaving enzyme 1 (BACE1), isoform A); (b) an engineered cleavage site;

XX CC and (c) a protease domain. (1) is capable of being cleaved at the

XX CC engineered cleavage site, and so releases a free protease domain that has

XX CC BACE1 activity. Also described: (1) a nucleic acid sequence encoding (1);

XX CC (2) a vector for expression of (1); and (3) a host cell expressing (1).

CC (I) has neuroprotective and nootropic activities, and can be used in gene  
CC therapy. (I) can be used for producing preparations of homogenously  
CC processed BACE1 that may be used for e.g. studying or treating diseases  
CC such as Alzheimer's disease or Down's syndrome. The human BACE1 gene is  
CC located on chromosome 11, more specifically to 11q23.2-23.3. The present  
CC sequence represents a mutant human BACE1 amino acid sequence, which is  
CC used in an example from the present invention.

XX Sequence 433 AA;

Query Match 100.0%; Score 2288; DB 8; Length 433;  
Best Local Similarity 100.0%; Pred. No. 6, 6e-216;  
Matches 433; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 TQHGIRLPRLSSGLGAPLGLFINETDEBEPEGRGSEVEMVDNLRGSGGGYVEMTV 60
D 1 TQHGIRLPRLSSGLGAPLGLFINETDEBEPEGRGSEVEMVDNLRGSGGGYVEMTV 60
QY 61 GSPPTNLIVDTGSSNFAYGAPHPFLHRYRORLSTYRDLRKGVVYPYTGKKEGEL 120
D 61 GSPPTNLIVDTGSSNFAYGAPHPFLHRYRORLSTYRDLRKGVVYPYTGKKEGEL 120
QY 121 GTDLVSIIPHGPNTVVRANIAAITESDKFINGSNWEGILGLAVAEIARPDLSLEPPFDSL 180
D 121 GTDLVSIIPHGPNTVVRANIAAITESDKFINGSNWEGILGLAVAEIARPDLSLEPPFDSL 180
QY 181 VKQTHVNLFSLOLCGAGFPPLNOSVLAIVGSMITIGIDHSLYTGSIMWTPYIRREWYE 240
D 181 VKQTHVNLFSLOLCGAGFPPLNOSVLAIVGSMITIGIDHSLYTGSIMWTPYIRREWYE 240
QY 241 VIIVRVEINGQDLKMDCKEYNDKSIYDSGTNLRPKKVEFAVAKSIKAASTTEKFPDG 300
D 241 VIIVRVEINGQDLKMDCKEYNDKSIYDSGTNLRPKKVEFAVAKSIKAASTTEKFPDG 300
QY 301 FMIAGEOLVCMQAGTTPWNIFPVISLYLMGEVTVNSFRITTLPOQYLRPVEDVATSQDDCY 360
D 301 FMIAGEOLVCMQAGTTPWNIFPVISLYLMGEVTVNSFRITTLPOQYLRPVEDVATSQDDCY 360
QY 361 KFAISQSSGTGTMGAVIMEGFYVFPDRARKRIGFAVSACHVDEFRITAAVSGPFTLME 420
D 361 KFAISQSSGTGTMGAVIMEGFYVFPDRARKRIGFAVSACHVDEFRITAAVSGPFTLME 420
QY 421 DCGYNIPQTDDEST 433
D 421 DCGYNIPQTDDEST 433
```

## RESULT 2

ADP83954  
ID ADP83954 standard; protein; 433 AA.

XX ADP83954;  
XX  
XX 23-SEP-2004 (first entry)

DE Human BACE1 mutant amino acid sequence SEQ ID NO:81.

XX human; beta-site amyloid precursor protein cleaving enzyme 1;  
KW beta-site APP cleaving enzyme 1; BACE1; BACE1 isoform A; chromosome 11;  
KW prodomain; engineered cleavage site; protease domain; neuroprotective;  
KW nootropic; gene therapy; Alzheimer's disease; Down's syndrome; mutant.

OS Homo sapiens.  
OS Synthetic.

XX WO2004056962-A2.

XX 08-JUL-2004.

PF 02-DEC-2003; 2003WO-US038314.

PR 04-DEC-2002; 2002US-0430984P.

XX

PA (SUNE-) SUNESIS PHARM INC.

XX Ballinger M;

DR WPI, 2004-507703/48.

PT New polypeptides for producing homogenously processed preparations of  
PT beta site amyloid precursor protein-cleaving enzyme comprises a  
PT prodomain, an engineered cleavage site and a protease domain.  
PS Example 1; SEQ ID NO 81, 40pp; English.

CC The present invention describes a polypeptide (I) comprising in order  
CC from the N-terminus to the C-terminus: (a) a prodomain comprising at  
CC least 6 contiguous amino acids of the 16 amino acid sequence of SEQ ID  
CC NO:3 (ADP83877), comprising residues 22-37 of SEQ ID NO:1 (ADP83876) which  
CC is the longest isoform of human beta-site amyloid precursor protein (APP)  
CC cleaving enzyme 1 (BACE1), isoform A); (b) an engineered cleavage site;  
CC and (c) a protease domain. (I) is capable of being cleaved at the  
CC engineered cleavage site, and so releases a free protease domain that has  
CC BACE1 activity. Also described: (1) a nucleic acid sequence encoding (I);  
CC (2) a vector for expression of (I); and (3) a host cell expressing (I).  
CC (I) has neuroprotective and nootropic activities, and can be used in gene  
CC therapy. (I) can be used for producing preparations of homogenously  
CC processed BACE1 that may be used for e.g. studying or treating diseases  
CC such as Alzheimer's disease or Down's syndrome. The human BACE1 gene is  
CC located on chromosome 11, more specifically to 11q23.2-23.3. The present  
CC sequence represents a mutant human BACE1 amino acid sequence, which is  
CC used in an example from the present invention.

XX Sequence 433 AA;

Query Match 99.9%; Score 2286; DB 8; Length 433;  
Best Local Similarity 99.8%; Pred. No. 1e-215;  
Matches 432; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 TQHGIRLPRLSSGLGAPLGLFINETDEBEPEGRGSEVEMVDNLRGSGGGYVEMTV 60
D 1 TQHGIRLPRLSSGLGAPLGLFINETDEBEPEGRGSEVEMVDNLRGSGGGYVEMTV 60
QY 61 GSPPTNLIVDTGSSNFAYGAPHPFLHRYRORLSTYRDLRKGVVYPYTGKKEGEL 120
D 61 GSPPTNLIVDTGSSNFAYGAPHPFLHRYRORLSTYRDLRKGVVYPYTGKKEGEL 120
QY 121 GTDLVSIIPHGPNTVVRANIAAITESDKFINGSNWEGILGLAVAEIARPDLSLEPPFDSL 180
D 121 GTDLVSIIPHGPNTVVRANIAAITESDKFINGSNWEGILGLAVAEIARPDLSLEPPFDSL 180
QY 181 VKQTHVNLFSLOLCGAGFPPLNOSVLAIVGSMITIGIDHSLYTGSIMWTPYIRREWYE 240
D 181 VKQTHVNLFSLOLCGAGFPPLNOSVLAIVGSMITIGIDHSLYTGSIMWTPYIRREWYE 240
QY 241 VIIVRVEINGQDLKMDCKEYNDKSIYDSGTNLRPKKVEFAVAKSIKAASTTEKFPDG 300
D 241 VIIVRVEINGQDLKMDCKEYNDKSIYDSGTNLRPKKVEFAVAKSIKAASTTEKFPDG 300
QY 301 FMIAGEOLVCMQAGTTPWNIFPVISLYLMGEVTVNSFRITTLPOQYLRPVEDVATSQDDCY 360
D 301 FMIAGEOLVCMQAGTTPWNIFPVISLYLMGEVTVNSFRITTLPOQYLRPVEDVATSQDDCY 360
QY 361 KFAISQSSGTGTMGAVIMEGFYVFPDRARKRIGFAVSACHVDEFRITAAVSGPFTLME 420
D 361 KFAISQSSGTGTMGAVIMEGFYVFPDRARKRIGFAVSACHVDEFRITAAVSGPFTLME 420
QY 421 DCGYNIPQTDDEST 433
D 421 DCGYNIPQTDDEST 433
```

## RESULT 3

ADP83951  
ID ADP83951 standard; protein; 433 AA.

XX



```

Db 61 GSPPTLTNIIVDTGSSNFAVGAAHPFLHRYYQRLSTYRDLRKGYVPTTQGMKGEGL 120
Qy 121 GTDLVSIHGHGPNVTVRANIAAITESDKFFINGSNWEGILGLAYAEIARPDLSLEPPDSL 180
Db 121 GTDLVSIHGHGPNVTVRANIAAITESDKFFINGSNWEGILGLAYAEIARPDLSLEPPDSL 180
Qy 181 VKQTHVPLFSLQCGAGFPLNOSBVLASVGGSMITGGIDHSLYTGSLSLWTPPIRREYYE 240
Db 181 VKQTHVPLFSLQCGAGFPLNOSBVLASVGGSMITGGIDHSLYTGSLSLWTPPIRREYYE 240
Qy 241 VIIVREINGQDLKMDCKEYNYDKSIYDSGTTNLRPKKVFEEAVKSIKKAASSTKFPDG 300
Db 241 VIIVREINGQDLKMDCKEYNYDKSIYDSGTTNLRPKKVFEEAVKSIKKAASSTKFPDG 300
Qy 301 FWLGEOLVCWQAGTTPNNIFPVISLYLMGEVTNQSFRITLLPQOYLRPVEDVATSODDCY 360
Db 301 FWLGEOLVCWQAGTTPNNIFPVISLYLMGEVTNQSFRITLLPQOYLRPVEDVATSODDCY 360
Qy 361 KFAISQSGTGVMGAVIMEGFYVFPDRARKRIGFAVSACHVHDEFRTAAVEGPFVTLDM 420
Db 361 KFAISQSGTGVMGAVIMEGFYVFPDRARKRIGFAVSACHVHDEFRTAAVEGPFVTLDM 420
Qy 421 DCGNINPQTDEST 433
Db 421 DCGNINPQTDEST 433

```

## RESULT 5

ABR61929 standard; protein; 456 AA.

ABR61929;  
12-SEP-2003 (first entry)  
Human promemapsin 2-T1 protein.

Memapsin 1; nootropic; neuroprotective; memapsin 2; beta secretase;  
beta-amyloid protein; Alzheimer's disease; promemapsin 2-T1; human.

Homo sapiens.

MO2003039454-A2.

15-MAY-2003.

23-OCT-2002; 2002WO-US034324.

23-OCT-2001; 2001US-0335952P.

27-NOV-2001; 2001US-0333545P.

14-JAN-2002; 2002US-0348646P.

20-JUN-2002; 2002US-0390804P.

19-JUL-2002; 2002US-0397557P.

19-JUL-2002; 2002US-0397619P.

(OKLA-) OKLAHOMA MEDICAL RES FOUND.

(UNIT ) UNIV ILLINOIS FOUN.

Thoen AK, Tang J, Bilcer G, Chang W, Hong L, Koelsch G, Loy J;

Turner RT;

WPI: 2003-541410/51.

N-PSDB; ACC84850.

New peptide compounds are memapsin beta secretase inhibitors used for

treating Alzheimer's disease.

Claim 96; Fig 11; 407pp; English.

CC to memapsin 1-beta secretase and reduce the accumulation of beta-amyloid  
CC protein. The compounds can be used for treating Alzheimer's disease. The  
CC present sequence represents a human promemapsin 2-T1 protein  
XX  
SQ Sequence 456 AA;

Query Match 99.1%; Score 2267; DB 6; Length 456;  
Best Local Similarity 99.1%; Pred. No. 8.4e-214;  
Matches 429; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

Qy 1 TQHGIRLRLRSLGAGAPLGLINLETDEPERGRGSPFVMDNLKSGSGGYVEMTV 60
Db 24 TQHGIRLRLRSLGAGAPLGLRPRETDEPERGRGSPFVMDNLKSGSGGYVEMTV 83
Qy 61 GSPPTLTNIIVDTGSSNFAVGAAHPFLHRYYQRLSTYRDLRKGYVPTTQGMKGEGL 120
Db 84 GSPPTLTNIIVDTGSSNFAVGAAHPFLHRYYQRLSTYRDLRKGYVPTTQGMKGEGL 143
Qy 121 GTDLVSIHGHGPNVTVRANIAAITESDKFFINGSNWEGILGLAYAEIARPDLSLEPPDSL 180
Db 144 GTDLVSIHGHGPNVTVRANIAAITESDKFFINGSNWEGILGLAYAEIARPDLSLEPPDSL 203
Qy 181 VKQTHVPLFSLQCGAGFPLNOSBVLASVGGSMITGGIDHSLYTGSLSLWTPPIRREYYE 240
Db 204 VKQTHVPLFSLQCGAGFPLNOSBVLASVGGSMITGGIDHSLYTGSLSLWTPPIRREYYE 263
Qy 241 VIIVREINGQDLKMDCKEYNYDKSIYDSGTTNLRPKKVFEEAVKSIKKAASSTKFPDG 300
Db 264 VIIVREINGQDLKMDCKEYNYDKSIYDSGTTNLRPKKVFEEAVKSIKKAASSTKFPDG 323
Qy 301 FWLGEOLVCWQAGTTPNNIFPVISLYLMGEVTNQSFRITLLPQOYLRPVEDVATSODDCY 360
Db 324 FWLGEOLVCWQAGTTPNNIFPVISLYLMGEVTNQSFRITLLPQOYLRPVEDVATSODDCY 383
Qy 361 KFAISQSGTGVMGAVIMEGFYVFPDRARKRIGFAVSACHVHDEFRTAAVEGPFVTLDM 420
Db 384 KFAISQSGTGVMGAVIMEGFYVFPDRARKRIGFAVSACHVHDEFRTAAVEGPFVTLDM 443
Qy 421 DCGNINPQTDEST 433
Db 444 DCGNINPQTDEST 456

```

## RESULT 6

AAB61334 standard; protein; 488 AA.

AAB61334;

02-APR-2001 (first entry)

Memapsin 2 protein.

Memapsin 2; catalytic; Alzheimer's.

Homo sapiens.

MO200100663-A2.

04-JAN-2001.

27-JUN-2000; 2000WO-US017661.

28-JUN-1999; 99US-0141363P.

30-NOV-1999; 99US-0168060P.

25-JAN-2000; 2000US-0177836P.

27-JAN-2000; 2000US-0178368P.

08-JUN-2000; 2000US-0210292P.

(OKLA-) OKLAHOMA MEDICAL RES FOUND.

Tang JUN, Lin X, Koelsch G;

XX



DR WPI; 2001-102885/11.  
XX Purified recombinant catalytically active memapsin 2, used to screen  
PT inhibitors of it, which are used to treat and prevent Alzheimer's  
PT disease.  
XX Claim 2; Page 73-75; 86pp; English.  
XX The present invention relates to a purified recombinant catalytically  
CC active memapsin 2. The invention may be used for isolating inhibitors  
CC which are used to treat or prevent Alzheimer's disease. The invention may  
CC also be used to screen for individuals more genetically prone to develop  
CC Alzheimer's disease  
XX  
SQ Sequence 488 AA;  
Query Match 99.1%; Score 2267; DB 4; Length 488;  
Best Local Similarity 99.1%; Pred. No. 9.3e-214;  
Matches 429; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
QY 1 TQHGIRLPRLRSGLGAPLGLLEINLETDEPEREPRGRGSEVEMVDNLRGSGGGYVEMTV 60  
DB 9 TQHGIRLPRLRSGLGAPLGLRLPRETDEPEREPRGRGSEVEMVDNLRGSGGGYVEMTV 68  
QY 61 GSPPTNLILVDTGSSNFAVGAAPPLHRYRQRLSSTYRDLRKGVVYPTQKMEGL 120  
DB 69 GSPPTNLILVDTGSSNFAVGAAPPLHRYRQRLSSTYRDLRKGVVYPTQKMEGL 128  
QY 121 GTDVLVSIPIHGPVNTVRANITAIETSDKPFINGSNMEGILGLAYAEIARPDSDLPEPFDL 180  
DB 129 GTDVLVSIPIHGPVNTVRANITAIETSDKPFINGSNMEGILGLAYAEIARPDSDLPEPFDL 188  
QY 181 VKQTHVPLNLFSLQCGAGPFLNQSSEVLASVGSMTIIGIDHSLYTGSLWYTPIRREMYE 240  
DB 189 VKQTHVPLNLFSLQCGAGPFLNQSSEVLASVGSMTIIGIDHSLYTGSLWYTPIRREMYE 248  
QY 241 VIIIRVEINGQDLKMDCKEYNDKSIYDSGTNLRLPKVFEAAVKSIIKAASSTEFKPDG 300  
DB 249 VIIIRVEINGQDLKMDCKEYNDKSIYDSGTNLRLPKVFEAAVKSIIKAASSTEFKPDG 308  
QY 301 FNLGEOLVCMQAGTTTPMNIIPVVISLYLMEVTNOSFRITILPQOYLRPVEDVATSODDCY 360  
DB 309 FNLGEOLVCMQAGTTTPMNIIPVVISLYLMEVTNOSFRITILPQOYLRPVEDVATSODDCY 368  
QY 361 KPAISQSSSTGTWGAIVMEGFYVFPDRARKRIGFAVSACHVDEFTAAVEGFVTLDMH 420  
DB 369 KPAISQSSSTGTWGAIVMEGFYVFPDRARKRIGFAVSACHVDEFTAAVEGFVTLDMH 428  
QY 421 DCGYNIPQTDST 433  
DB 429 DCGYNIPQTDST 441  
RESULT 7  
AAB6572  
ID AAB6572 standard; protein; 488 AA.  
XX  
AC AAB6572;  
XX  
DT 12-APR-2001 (first entry)  
XX  
DE Human memapsin 2.  
XX  
KW Human; memapsin 2; neurotrophic; neuroprotective; amyloid precursor protein;  
KW APP; memapsin 2 inhibitor; Alzheimer's disease.  
XX  
XX Homo sapiens.  
XX  
XX OS  
XX PN WO20010665-A2.  
XX  
XX 04-JAN-2001.  
XX  
XX PD  
XX 27-JUN-2000; 2000WO-US017742.  
PF

XX  
PR 28-JUN-1999; 99US-0141363P.  
PR 30-NOV-1999; 99US-0168060P.  
PR 25-JAN-2000; 2000US-0177836P.  
PR 27-JAN-2000; 2000US-0178368P.  
PR 08-JUN-2000; 2000US-0210292P.  
XX  
PA (OKLA-) OKLAHOMA MEDICAL RES FOUND.  
PA (UNIT ) UNITV ILLINOIS FOUND.  
XX  
PI Tang JUN, Hong L, Ghosh AK;  
XX  
DR WPI: 2001-137933/14.  
DR N-PSDB; AAF31848.  
XX  
PT Novel memapsin 2 inhibitors which bind to active site of memapsin 2  
PT having 2 catalytic aspartic residues and substrate binding cleft, used to  
PT treat Alzheimer's disease by blocking amyloid precursor protein cleavage.  
PS Example 1; Page 72-74; 86pp; English.  
XX  
XX The present sequence is given in a specification relating to an inhibitor  
CC of catalytically active memapsin 2. The inhibitor binds to the memapsin 2  
CC active site, which is defined by the presence of two catalytic aspartic  
CC residues and a substrate binding cleft. The inhibitor is useful for the  
CC treatment and diagnosis of Alzheimer's disease. It is useful in screens  
CC for individuals with a genetic predisposition to Alzheimer's disease. The  
CC inhibitor is useful as a reagent for specifically binding to memapsin 2  
CC or memapsin 2 analogues and for aiding in memapsin 2 isolation,  
CC purification and characterisation  
XX  
SQ Sequence 488 AA;  
Query Match 99.1%; Score 2267; DB 4; Length 488;  
Best Local Similarity 99.1%; Pred. No. 9.3e-214;  
Matches 429; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
QY 1 TQHGIRLPRLRSGLGAPLGLLEINLETDEPEREPRGRGSEVEMVDNLRGSGGGYVEMTV 60  
DB 9 TQHGIRLPRLRSGLGAPLGLRLPRETDEPEREPRGRGSEVEMVDNLRGSGGGYVEMTV 68  
QY 61 GSPPTNLILVDTGSSNFAVGAAPPLHRYRQRLSSTYRDLRKGVVYPTQKMEGL 120  
DB 69 GSPPTNLILVDTGSSNFAVGAAPPLHRYRQRLSSTYRDLRKGVVYPTQKMEGL 128  
QY 121 GTDVLVSIPIHGPVNTVRANITAIETSDKPFINGSNMEGILGLAYAEIARPDSDLPEPFDL 180  
DB 129 GTDVLVSIPIHGPVNTVRANITAIETSDKPFINGSNMEGILGLAYAEIARPDSDLPEPFDL 188  
QY 181 VKQTHVPLNLFSLQCGAGPFLNQSSEVLASVGSMTIIGIDHSLYTGSLWYTPIRREMYE 240  
DB 189 VKQTHVPLNLFSLQCGAGPFLNQSSEVLASVGSMTIIGIDHSLYTGSLWYTPIRREMYE 248  
QY 241 VIIIRVEINGQDLKMDCKEYNDKSIYDSGTNLRLPKVFEAAVKSIIKAASSTEFKPDG 300  
DB 249 VIIIRVEINGQDLKMDCKEYNDKSIYDSGTNLRLPKVFEAAVKSIIKAASSTEFKPDG 308  
QY 301 FNLGEOLVCMQAGTTTPMNIIPVVISLYLMEVTNOSFRITILPQOYLRPVEDVATSODDCY 360  
DB 309 FNLGEOLVCMQAGTTTPMNIIPVVISLYLMEVTNOSFRITILPQOYLRPVEDVATSODDCY 368  
QY 361 KPAISQSSSTGTWGAIVMEGFYVFPDRARKRIGFAVSACHVDEFTAAVEGFVTLDMH 420  
DB 369 KPAISQSSSTGTWGAIVMEGFYVFPDRARKRIGFAVSACHVDEFTAAVEGFVTLDMH 428  
QY 421 DCGYNIPQTDST 433  
DB 429 DCGYNIPQTDST 441  
RESULT 8  
ABG78372  
ID ABG78372 standard; protein; 488 AA.  
PF

```

XX AC ABG78372;
XX XX
XX DT 15-NOV-2002 (first entry)
XX DE Human promemapsin 2.
XX KW Human; enzyme; memapsin 2; aspartic protease; beta secretase;
XX KW degenerative disease; Alzheimer's disease; amyloid precursor protein;
XX KW APP; neuroprotective; nootropic; inhibitor;
XX OS substrate side-chain preference.
XX OS Homo sapiens.
XX PN WO200253594-A2.
XX PD 11-JUL-2002.
XX XX
XX PF 28-DEC-2001; 2001WO-US050826.
XX PR 28-DEC-2000; 2000US-0258705P.
XX PR 14-MAR-2001; 2001US-0275756P.
XX XX
XX PA (OKLA-) OKLAHOMA MEDICAL RES FOUND.
XX PA (UNIT ) UNIT ILLINOIS FOUND.
XX PI Tang JUN, Koelsch G, Ghosh AK;
XX DR WPI; 2002-619088/66.
XX PT New memapsin 2 activity inhibitor useful in treatment of e.g. Alzheimer's
XX PT disease.
XX PS Claim 7; Fig 7; 74pp; English.
XX XX
CC The invention relates to an inhibitor of catalytically active memapsin 2
CC (an aspartic protease which can cleave at beta secretase sites), which
CC binds to the active site of memapsin 2 defined by the presence of two
CC catalytic aspartic residues and substrate binding cleft. Also included is
CC a method of determination of the substrate side-chain preference in
CC memapsin 2 sub-sites comprising: (a) reacting a mixture of memapsin 2
CC substrates with memapsin 2, and determining the sub-site preference of
CC memapsin 2 by determining relative initial hydrolysis rates of the
CC mixture of memapsin 2 substrates; or (b) preparing a combinatorial
CC library of memapsin 2 inhibitors containing a base sequence taken from
CC OM99-2 (Glu-Val-Asn-Leu-Ala-Glu-phe), probing the library of
CC inhibitors with memapsin 2 which binds to several inhibitors to generate
CC several bound memapsin 2, and detecting the bound memapsin 2 with an
CC antibody raised to memapsin 2 and an alkaline phosphatase conjugated
CC secondary antibody. The inhibitors may be used in the manufacture of a
CC medicament for the treatment of Alzheimer's disease since memapsin 2 may
CC be involved in the cleavage of amyloid precursor protein (APP), and for
CC determining the substrate side-chain preference in memapsin 2 sub-sites.
CC The present sequence represents human memapsin 2 (either prepromemapsin 2
CC or mature memapsin)
XX
SQ Sequence 488 AA;
Query Match 99.1%; Score 2267; DB 5; Length 488;
Best Local Similarity 99.1%; Pred. No. 9,3e-214;
Matches 429; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 TGHGRLPLRSGLGAPLGLTINLETDEPEEPGRGSGFVEMVNDLRKSGQGYVENVTV 60
DB 9 TGHGRLPLRSGLGAPLGLTINLETDEPEEPGRGSGFVEMVNDLRKSGQGYVENVTV 68
QY 61 GSPPTQNTLVDTGSSNFVGAAPFLHRYRQRLSTVYDLKRGVVPPTQKMGCEL 120
DB 69 GSPPTQNTLVDTGSSNFVGAAPFLHRYRQRLSTVYDLKRGVVPPTQKMGCEL 128
QY 121 GTDLVSIHGPVNTVRRANIAAITESDKPFINGSMWEGITGLAVYELARPDDSLPFPDSL 180
DB 129 GTDLVSIHGPVNTVRRANIAAITESDKPFINGSMWEGITGLAVYELARPDDSLPFPDSL 188

```

```

QY 181 VKQTHVPLNLSQLCGAGPLNQSEVLASVGSMTIGSIDHSLYTGSLMTPRIRREWYE 240
DB 189 VKQTHVPLNLSQLCGAGPLNQSEVLASVGSMTIGSIDHSLYTGSLMTPRIRREWYE 248
QY 241 VIVRVEINQDILKNDCKEYNDKSIYDSGTTNLRPKQVEAAVKSIIKAASTEKEPDG 300
DB 249 VIVRVEINQDILKNDCKEYNDKSIYDSGTTNLRPKQVEAAVKSIIKAASTEKEPDG 308
QY 301 FHLGEOLVQWQAGTTPMNIFFVISLYLMGEVYTNOSFRITLLPOQYLAPVEVVAISODDCY 360
DB 309 FHLGEOLVQWQAGTTPMNIFFVISLYLMGEVYTNOSFRITLLPOQYLAPVEVVAISODDCY 368
QY 361 KFAISQSSGTVMGAVIMEGFYVFPDRARKRIGFAVACVHDFRTAAVEGPFVTLDME 420
DB 369 KFAISQSSGTVMGAVIMEGFYVFPDRARKRIGFAVACVHDFRTAAVEGPFVTLDME 428
QY 421 DCGYNIPTQDST 433
DB 429 DCGYNIPTQDST 441

```

## RESULT 9

```

AAU99488
ID AAU99488 standard; protein; 488 AA.
XX
XX AAU99488;
XX
XX 07-OCT-2002 (first entry)
XX
XX Human memapsin 2.
XX
XX
XX Human; memapsin 2; beta secretase; aspartic protease; APP;
XX KW beta-amyloid precursor protein; amyloid plaque; Alzheimer's disease;
XX KW neuroprotective; nootropic; enzyme.
XX OS Homo sapiens.
XX
XX US2002049303-A1.
XX
XX 25-APR-2002.
XX
XX 28-FEB-2001; 2001US-00796264.
XX
XX 28-JUN-1999; 99US-0141363P.
XX PR 30-NOV-1999; 99US-0168060P.
XX PR 25-JAN-2000; 2000US-0177836P.
XX PR 27-JUN-2000; 2000US-0178368P.
XX PR 27-JUN-2000; 2000US-00604608.
XX
XX (TANG/) TANG J J N.
XX PA (LINK/) LIN X.
XX PA (KOEL/) KOELSCH G.
XX PA (HONG/) HONG L.
XX
XX Tang JUN, Lin X, Koelsch G, Hong L;
XX
XX WPI; 2002-507280/54.
XX DR N-PSDB; ABX88641.
XX
XX New recombinant catalytically active memapsin 2, useful to screen for
XX PT inhibitors of memapsin 2 which can be used to prevent and treat
XX PT Alzheimer's disease.
XX PS Claim 2; Page 22-23; 44pp; English.
XX
XX The present invention relates to methods for the production of purified,
XX CC recombinant catalytically active, memapsin 2 (beta secretase). Memapsin
XX CC 2, a member of the aspartic protease family, cleaves beta-amyloid
XX CC precursor protein (APP) found in amyloid plaques. The recombinant
XX CC memapsin 2 is useful for identifying inhibitors of memapsin 2 in the
XX CC design of drugs for the treatment and/or prevention of Alzheimer's
XX CC disease. The recombinant memapsin 2 can be used to immunise against

```

4.  
CC Alzheimer's disease. The present sequence represents human memapsin 2  
XX  
SQ Sequence 488 AA;

Query Match 99.1%; Score 2267; DB 5; Length 488;  
Best Local Similarity 99.1%; Pred. No. 9, 3e-214;  
Matches 429; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 TQHGIRLPRLRSGLGAPLGLINLETDEBPBGRGRSVFVENVDLRGSGGGYVEMTV 60  
DB 9 TQHGIRLPRLRSGLGAPLGLRLPRETDEBPBGRGRSVFVENVDLRGSGGGYVEMTV 68  
QY 61 GSPPTLNTLVDTGSSNFVGAAPHPFLHRYTORLSSTYRDLRKGVVPTQGWKEGL 120  
DB 69 GSPPTLNTLVDTGSSNFVGAAPHPFLHRYTORLSSTYRDLRKGVVPTQGWKEGL 128  
QY 121 GTDLVSIPIPGPNVTVRANIAITTESDKFFINSNMEGILGLAYAEIARPDLSLEPPFDSL 180  
DB 129 GTDLVSIPIPGPNVTVRANIAITTESDKFFINSNMEGILGLAYAEIARPDLSLEPPFDSL 188  
QY 181 VQOTHVPNLFSIQLCGAGFPLNQSEVLASVGSMTIIGIDHSLYTGSMTYTPIRREMYE 240  
DB 189 VQOTHVPNLFSIQLCGAGFPLNQSEVLASVGSMTIIGIDHSLYTGSMTYTPIRREMYE 248  
QY 241 VTIIRVEINGODLKMDCKEYNTDKSIVDSGTTNLRPKVYFEAAVKSITAASTTEKFPDG 300  
DB 249 VTIIRVEINGODLKMDCKEYNTDKSIVDSGTTNLRPKVYFEAAVKSITAASTTEKFPDG 308  
QY 301 FMLGEOLVCMQAGTTTWNIFPVISLYLMGEVNTQSRITILPQOYLRPVEDVATSGDDCY 360  
DB 309 FMLGEOLVCMQAGTTTWNIFPVISLYLMGEVNTQSRITILPQOYLRPVEDVATSGDDCY 368  
QY 361 KFAISQSTGTGTVGAVIMGEFVYVFDRAKRIIGFAVSACHVDEFTAAVEGPFVTLDM 420  
DB 369 KFAISQSTGTGTVGAVIMGEFVYVFDRAKRIIGFAVSACHVDEFTAAVEGPFVTLDM 428  
QY 421 DCGYNIPQTDST 433  
DB 429 DCGYNIPQTDST 441

## RESULT 10

ABG76101  
ID ABG76101 standard; protein; 488 AA.

XX  
AC ABG76101;

DT 01-MAY-2003 (first entry)

XX Human partial memapsin 2.

XX Human; memapsin 2; beta-secretase; beta-amyloid precursor protein;  
KW beta-amyloid peptide; Alzheimer's disease; neuroprotective;  
KW enzyme.

XX Homo sapiens.

OS  
XX  
PN US2002164760-A1.

XX  
PD 07-NOV-2002.

XX  
PF 28-FEB-2001; 2001US-00795903.

XX  
PR 28-JUN-1999; 99US-0141363P.

XX  
PR 30-NOV-1999; 99US-0168060P.

XX  
PR 25-JAN-2000; 2000US-0177836P.

XX  
PR 27-JAN-2000; 2000US-0178368P.

XX  
PR 08-JUN-2000; 2000US-0210292P.

XX  
PR 27-JUN-2000; 2000US-00604608.

XX  
PA (OKLA-) OKLAHOMA MEDICAL RES FOUND.  
XX  
PI Lin X, Koelsch G, Tang JUN;

XX  
DR MPI: 2003-255218/25.  
DR N-ESDB; ABX11591.

PT New purified recombinant catalytically active memapsin 2 (beta-  
secretase), useful for designing and screening of specific inhibitors for  
the diagnosis, prevention and/or treatment of Alzheimer's disease.

XX  
PS Claim 2; Page 22-23; 44pp; English.

XX  
CC The invention relates to a purified recombinant catalytically active  
CC memapsin 2, a beta-secretase which produces the beta-amyloid peptide from  
CC the beta amyloid precursor protein. Also included are producing the above  
CC memapsin 2 (comprising refolding the recombinant memapsin 2 under  
CC conditions which dissociate and then slowly refold the enzyme into a  
CC catalytically active form), isolating inhibitors of cleavage by memapsin  
CC 2 (comprising adding to one or more potential inhibitors the memapsin 2  
CC and a substrate for memapsin 2 and screening for decreased cleavage of  
CC the substrate by the inhibitors), designing or obtaining inhibitors of  
CC the memapsin 2 (comprising modelling an inhibitor based on the  
CC crystallisation coordinates of memapsin 2 or the parameters given in the  
CC specification), a database comprising binding properties and chemical  
CC structures of compounds designed or screened by the method above and  
CC treating or preventing Alzheimer's disease (comprising administering to a  
CC patient an inhibitor of memapsin 2 which binds to the active site of the  
CC memapsin 2 defined by the presence of 2 catalytic aspartic residues and  
CC substrate binding cleft or immunising an individual with the above  
CC memapsin 2 to elicit an amount of antibodies to reduce the cleavage by  
CC endogenous memapsin 2). The memapsin 2 is useful in designing and  
CC screening of specific inhibitors for the diagnosis, prevention and/or  
CC treatment of Alzheimer's disease. The present sequence represents  
CC memapsin 2

XX  
SQ Sequence 488 AA;

Query Match 99.1%; Score 2267; DB 6; Length 488;  
Best Local Similarity 99.1%; Pred. No. 9, 3e-214;  
Matches 429; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 TQHGIRLPRLRSGLGAPLGLINLETDEBPBGRGRSVFVENVDLRGSGGGYVEMTV 60  
DB 9 TQHGIRLPRLRSGLGAPLGLRLPRETDEBPBGRGRSVFVENVDLRGSGGGYVEMTV 68  
QY 61 GSPPTLNTLVDTGSSNFVGAAPHPFLHRYTORLSSTYRDLRKGVVPTQGWKEGL 120  
DB 69 GSPPTLNTLVDTGSSNFVGAAPHPFLHRYTORLSSTYRDLRKGVVPTQGWKEGL 128  
QY 121 GTDLVSIPIPGPNVTVRANIAITTESDKFFINSNMEGILGLAYAEIARPDLSLEPPFDSL 180  
DB 129 GTDLVSIPIPGPNVTVRANIAITTESDKFFINSNMEGILGLAYAEIARPDLSLEPPFDSL 188  
QY 181 VQOTHVPNLFSIQLCGAGFPLNQSEVLASVGSMTIIGIDHSLYTGSMTYTPIRREMYE 240  
DB 189 VQOTHVPNLFSIQLCGAGFPLNQSEVLASVGSMTIIGIDHSLYTGSMTYTPIRREMYE 248  
QY 241 VTIIRVEINGODLKMDCKEYNTDKSIVDSGTTNLRPKVYFEAAVKSITAASTTEKFPDG 300  
DB 249 VTIIRVEINGODLKMDCKEYNTDKSIVDSGTTNLRPKVYFEAAVKSITAASTTEKFPDG 308  
QY 301 FMLGEOLVCMQAGTTTWNIFPVISLYLMGEVNTQSRITILPQOYLRPVEDVATSGDDCY 360  
DB 309 FMLGEOLVCMQAGTTTWNIFPVISLYLMGEVNTQSRITILPQOYLRPVEDVATSGDDCY 368  
QY 361 KFAISQSTGTGTVGAVIMGEFVYVFDRAKRIIGFAVSACHVDEFTAAVEGPFVTLDM 420  
DB 369 KFAISQSTGTGTVGAVIMGEFVYVFDRAKRIIGFAVSACHVDEFTAAVEGPFVTLDM 428  
QY 421 DCGYNIPQTDST 433  
DB 429 DCGYNIPQTDST 441

## RESULT 11

ADA74817	ID	ADA74817 standard; protein; 488 AA.
ADA74817	AC	ADA74817;
20-NOV-2003	DT	(first entry)
Human memapsin 2 protein.	DE	
memapsin 2; beta-secretase; amyloid precursor protein; APP; aspartic proteinase 2; ASP2; neurotrophic; neuroprotective; Alzheimer's disease; human; enzyme.	KW	
Homo sapiens.	OS	
Key	FT	Location/Qualifiers
Region	FT	1..180
Region	FT	/note= "N-terminal lobe"
Region	FT	181..385
Region	FT	/note= "C-terminal lobe"
Region	FT	359..393
Region	FT	/note= "C-terminal extension"
Region	FT	369..376
Region	FT	/note= "Beta-sheet 1"
Region	FT	378..383
Region	FT	/note= "Alpha-helix 1"
US6545127-B1.	PN	
08-APR-2003.	XX	
27-JUN-2000; 2000US-00604608.	XX	
28-JUN-1999; 99US-0141363P.	PR	
30-NOV-1999; 99US-0168060P.	PR	
25-JAN-2000; 2000US-0177836P.	PR	
27-JAN-2000; 2000US-0178368P.	PR	
08-JUN-2000; 2000US-0210292P.	PR	
(OKLA-) OKLAHOMA MEDICAL RES FOUND.	PA	
Tang JUN, Lin X, Koelsch G, Hong L;	PI	
WPI; 2003-566587/53.	DR	
N-PSDB; ADA74816.	XX	
Novel memapsin 2 protein that cleaves a beta-secretase site of an amyloid precursor protein is useful in the design and screening of specific inhibitors for treating and preventing Alzheimer's disease.	PT	
Disclosure; Col 39-42; 44pp; English.	XX	
The invention relates to a novel method which comprises the production of a purified, catalytically active, recombinant memapsin 2 (beta-secretase) protein where the memapsin protein is expressed in a bacterial cell and cleaves the beta-secretase site of an amyloid precursor protein (APP).	CC	
Memapsin 2, also known as aspartic proteinase 2 (ASP2), belongs to the aspartic protease family and demonstrates neurotrophic and neuroprotective activities. The protein of the invention may be useful in the design and screening of specific inhibitors which are useful in treating and preventing Alzheimer's disease. The current sequence is that of the human memapsin 2 protein of the invention.	CC	
Sequence 488 AA;	XX	

Query Match 99.1%; Score 2267; DB 6; Length 488;  
 Best Local Similarity 99.1%; Pred. No. 9.3e-214;  
 Matches 429; Conservative 1; Mismatches 3; Indels 0; Gaps 0

```
OY      1 TQHGIRLPLRSGGLGAPLGLINLETDEBPPEPGRGSFVEMVDNLRGSKSGGYVENTV   60
          ||||| : |||||
Db       9 TQHGIRLPLRSGGLGAPLGLRLPRETDEBPPEPGRGSFVEMVDNLRGSKSGGYVENTV   68
```

QY	61	GSPQOTNIIIVDTGSSNFAVGAAPHPFLRRYYQROLSSTYDRLKRGVYPTTGKMGEL	120
Db	69	GSPQOTNIIIVDTGSSNFAVGAAPHPFLRRYYQROLSSTYDRLKRGVYPTTGKMGEL	128
QY	121	GTDLVSIIPHGPNTVBARNIAITESPCKPFIINGSNMEGLGLAYAEIARPDLSLEPPFDSL	180
Db	129	GTDLVSIIPHGPNTVBARNIAITESPCKPFIINGSNMEGLGLAYAEIARPDLSLEPPFDSL	188
QY	181	VKQTHVENLFSLQLCGAGFPINQSEVLASVGGSMIIIGSIDHSLTGSIMWTPPIRREWYE	240
Db	189	VKQTHVENLFSLQLCGAGFPINQSEVLASVGGSMIIIGSIDHSLTGSIMWTPPIRREWYE	248
QY	241	VIIYRVINQDILKMKCKEYNTDKSIYDSGTTNLRPKYFEAAVKSIKAASSTKFPDG	300
Db	249	VIIYRVINQDILKMKCKEYNTDKSIYDSGTTNLRPKYFEAAVKSIKAASSTKFPDG	308
QY	301	FMLGEOLVCMQAGTPEPMNIFPVISLYLIMGVLTNOSFRITLLPQOYLAPVVDVATSDDDC	360
Db	309	FMLGEOLVCMQAGTPEPMNIFPVISLYLIMGVLTNOSFRITLLPQOYLAPVVDVATSDDDC	368
QY	361	KFAISQSTGTGVAIVIMEGFYVVDPRARKIGFAVSACHVDEFPRTAAVEGFPVILDM	420
Db	369	KFAISQSTGTGVAIVIMEGFYVVDPRARKIGFAVSACHVDEFPRTAAVEGFPVILDM	428
QY	421	DCGYNIPQTDEST 433	
Db	429	DCGYNIPQTDEST 441	

RESULT 12  
AA94767  
ID AA94767 standard; protein; 501 AA

DT 12-FEB-2001 (first entry)

Human beta-secretase amino acid sequence.

KM Beta-secretase; enzyme; amyloid plaque; Alzheimer's disease; human,  
Down's syndrome; amyloid angiopathy; gene therapy; neuroprotective,  
...

OS Homo sapiens.

EH	Key	Location/Qualifiers
FT	Peptide	1..45
FT		/label= putative signal peptide
FT	Protein	46..501
FT		/label= Beta-secretase

PN WO200058479-A1

PD 05-OCT-2000.

PF 23-MAR-2000; 2000WO-US007755.

PR 26-MAR-1999; 99US-00277229.

PA (AMGE-) AMGEN INC.

PI Citron M, Vassar RJ, Bennett BD;

DR WPI; 2000-594643/56.

XX

**PT** Isolated beta-secretase nucleic acids and encoded polypeptides, useful for diagnosis and gene therapy of Alzheimer's disease.

PS Claim 1; Fig 4; 145pp; English.

CC This invention relates to 3 nucleotide sequences encoding beta-secretase  
CC proteins. Beta-secretase is an enzyme involved in the production of one  
CC of the components of amyloid plaques involved in Alzheimer's disease. The

invention includes an expression vector comprising the nucleotide sequence, a host cell comprising the expression vector, and a process for producing the protein through culturing the transformed cells. Also included in the invention are a polypeptide derivative of the beta-secretase protein, a fusion protein comprising beta-secretase fused to a heterologous amino acid sequence, and a method for modulating the levels of beta-secretase polypeptide in a mammal comprising administering the polynucleotide sequence. Beta-secretase exhibits neuroprotective and neurotrophic activity. The beta-secretase nucleotide sequence may be used to map locations of the beta-secretase gene and related genes on chromosomes and as hybridization probes in diagnostic assays to test for the presence of beta-secretase DNA or RNA, such as in Alzheimer's disease, Down's syndrome, and amyloid angiopathy. The nucleotide sequence may also be used as anti-sense inhibitors of beta-secretase expression, in gene therapy of Alzheimer's disease, and for the identification of compounds that modulate beta-secretase activity. Antibodies to the beta-secretase protein may be used for in vitro and in vivo diagnostic purposes to detect the presence of beta-secretase polypeptide in a body fluid or cell sample. The present sequence represents the human beta-secretase protein

Sequence 501 AA;

Query Match 99.1%; Score 2267; DB 3; Length 501;  
Best Local Similarity 99.1%; Pred. No. 9.7e-214;  
Matches 429; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

1 TGHGRLPLRSGLGAPLGLLEINLETDSEBEPGRGRSFVEMVDNLRGSGGQYVEMTV 60  
22 TGHGRLPLRSGLGAPLGLRLPRETDEBEPGRGRSFVEMVDNLRGSGGQYVEMTV 81  
61 GSPPTLNLIVDTGSSNFVGAAPHPFLHRYRQQLSSTRDLRKGVVYPYTGKKEGEL 120  
82 GSPPTLNLIVDTGSSNFVGAAPHPFLHRYRQQLSSTRDLRKGVVYPYTGKKEGEL 141  
121 GTDLVSIPIGPNVTYVANIATIESDKPFINSNMGGIIGLAAEIAARPDDSLPEPFDSL 180  
142 GTDLVSIPIGPNVTYVANIATIESDKPFINSNMGGIIGLAAEIAARPDDSLPEPFDSL 201  
181 VKQTHVPNLFSIQLCGAGPFLNOSVLAIVSGSMIIGIDHSLYTSLWTPPIRRMYEY 240  
202 VKQTHVPNLFSIQLCGAGPFLNOSVLAIVSGSMIIGIDHSLYTSLWTPPIRRMYEY 261  
241 VIIVRVEINGODLKMDCKEYNDKSIYDSGTTNLRPKVFEAAVYSIKAASTTEKFPDG 300  
262 VIIVRVEINGODLKMDCKEYNDKSIYDSGTTNLRPKVFEAAVYSIKAASTTEKFPDG 321  
301 FWLGEOLVCMQAGTTTPMNIIPVVISLYLMGEVTVNOSFRITILLPOQYLRPEVDVATSDDDCY 360  
322 FWLGEOLVCMQAGTTTPMNIIPVVISLYLMGEVTVNOSFRITILLPOQYLRPEVDVATSDDDCY 381  
361 KFAISOSSTGTWGAIVMEGFYVFPDRARRKIGFAVSACHVHDEFRTAAVEGFPVTLDM 420  
382 KFAISOSSTGTWGAIVMEGFYVFPDRARRKIGFAVSACHVHDEFRTAAVEGFPVTLDM 441  
421 DCGYNIPQTDST 433  
442 DCGYNIPQTDST 454

RESULT 13  
AAB07896  
ID AAB07896 standard; protein; 501 AA.  
XX  
AC AAB07896;  
XX  
DT 14-NOV-2000 (first entry)  
XX  
XX Amino acid sequence of a human beta-secretase enzyme.  
DE Beta-secretase; beta-amyloid precursor protein; beta-amyloid peptide;  
KM amyloid plaque component; Alzheimer's disease; amyloidogenic disease;  
XX inhibitor.

OS Homo sapiens.  
XX  
XX WO2000047618-A2.  
XX  
XX 17-AUG-2000.  
XX  
XX 10-FEB-2000; 2000WO-US003819.  
XX  
XX 10-FEB-1999; 99US-0119571P.  
XX 15-JUN-1999; 99US-0139172P.  
XX  
XX (ELAN-) ELAN PHARM INC.  
XX  
XX Anderson UP, Baai G, Doane MT, Frigon N, John V, Power M,  
XX Sima S, Tatsuno G, Tung J, Wang S, Mcconlogue L,  
XX WPI; 2000-533011/48.  
XX N-PSDB; AAM59550, AAM59551.  
XX  
XX Purified beta-secretase protein used in assays to discover inhibitors  
XX which can be used for the treatment of amyloidogenic diseases e.g.  
XX Alzheimer's disease.  
XX  
XX Claim 17; Fig 2A; 121p; English.

The specification describes a beta-secretase enzyme. The enzyme cleaves beta-amyloid precursor protein to produce beta-amyloid peptide. This enzyme is therefore implicated in the production of amyloid plaque components which accumulate in the brains of individuals afflicted with Alzheimer's disease. Inhibitors of beta-secretase are administered to a mammalian subject e.g. with Alzheimer's disease or Alzheimer's disease-like pathology to test if they maintain or improve cognitive ability or reduce the plaque burden. The compounds are used for the treatment of amyloidogenic diseases e.g. Alzheimer's disease. The present sequence represents a human beta-secretase enzyme

Sequence 501 AA;

Query Match 99.1%; Score 2267; DB 3; Length 501;  
Best Local Similarity 99.1%; Pred. No. 9.7e-214;  
Matches 429; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

1 TGHGRLPLRSGLGAPLGLLEINLETDSEBEPGRGRSFVEMVDNLRGSGGQYVEMTV 60  
22 TGHGRLPLRSGLGAPLGLRLPRETDEBEPGRGRSFVEMVDNLRGSGGQYVEMTV 81  
61 GSPPTLNLIVDTGSSNFVGAAPHPFLHRYRQQLSSTRDLRKGVVYPYTGKKEGEL 120  
82 GSPPTLNLIVDTGSSNFVGAAPHPFLHRYRQQLSSTRDLRKGVVYPYTGKKEGEL 141  
121 GTDLVSIPIGPNVTYVANIATIESDKPFINSNMGGIIGLAAEIAARPDDSLPEPFDSL 180  
142 GTDLVSIPIGPNVTYVANIATIESDKPFINSNMGGIIGLAAEIAARPDDSLPEPFDSL 201  
181 VKQTHVPNLFSIQLCGAGPFLNOSVLAIVSGSMIIGIDHSLYTSLWTPPIRRMYEY 240  
202 VKQTHVPNLFSIQLCGAGPFLNOSVLAIVSGSMIIGIDHSLYTSLWTPPIRRMYEY 261  
241 VIIVRVEINGODLKMDCKEYNDKSIYDSGTTNLRPKVFEAAVYSIKAASTTEKFPDG 300  
262 VIIVRVEINGODLKMDCKEYNDKSIYDSGTTNLRPKVFEAAVYSIKAASTTEKFPDG 321  
301 FWLGEOLVCMQAGTTTPMNIIPVVISLYLMGEVTVNOSFRITILLPOQYLRPEVDVATSDDDCY 360  
322 FWLGEOLVCMQAGTTTPMNIIPVVISLYLMGEVTVNOSFRITILLPOQYLRPEVDVATSDDDCY 381  
361 KFAISOSSTGTWGAIVMEGFYVFPDRARRKIGFAVSACHVHDEFRTAAVEGFPVTLDM 420  
382 KFAISOSSTGTWGAIVMEGFYVFPDRARRKIGFAVSACHVHDEFRTAAVEGFPVTLDM 441  
421 DCGYNIPQTDST 433  
442 DCGYNIPQTDST 454

```

RESULT 14
ABG78374
ID   ABG78374 standard; protein; 501 AA.
XX
AC   ABG78374;
XX
DT   15-NOV-2002 (first entry)
XX
DE   Human prepromemapsin 2.
XX
KW   Human; enzyme; memapsin 2; aspartic protease; beta secretase;
KW   degenerative disease; Alzheimer's disease; amyloid precursor protein;
KW   APP; neuroprotective; nootropic; inhibitor;
KW   substrate side-chain preference.
XX
OS   Homo sapiens.
XX
PN   WO200253594-A2.
XX
PD   11-JUL-2002.
XX
PF   28-DEC-2001; 2001WO-US050826.
XX
PR   28-DEC-2000; 2000US-0258705P.
XX
PR   14-MAR-2001; 2001US-0275756P.
XX
PA   (OKLA-) OKLAHOMA MEDICAL RES FOUND.
XX
PI   (UNIT ) UNIV ILLINOIS FOUND.
XX
PT   Tang JUN, Koelech G, Ghosh AK;
XX
DR   WPI; 2002-619088/66.
XX
PT   New memapsin 2 activity inhibitor useful in treatment of e.g. Alzheimer's
XX   disease.
XX
PS   Disclosure, Fig 9; 74pp; English.
XX
CC   The invention relates to an inhibitor of catalytically active memapsin 2
CC   (an aspartic protease which can cleave at beta secretase sites), which
CC   binds to the active site of memapsin 2 defined by the presence of two
CC   catalytic aspartic residues and substrate binding cleft. Also included is
CC   a method of determination of the substrate side-chain preference in
CC   memapsin 2 sub-sites comprising: (a) reacting a mixture of memapsin 2
CC   substrates with memapsin 2, and determining the sub-site preference of
CC   memapsin 2 by determining relative initial hydrolysis rates of the
CC   mixture of memapsin 2 substrates; or (b) preparing a combinatorial
CC   library of memapsin 2 inhibitors containing a base sequence taken from
CC   OM99-2 (Glu-Val-Leu-Ala-Ala-Glu-phe), probing the library of
CC   inhibitors with memapsin 2 which binds to several inhibitors to generate
CC   several bound memapsin 2, and detecting the bound inhibitors 2 with an
CC   antibody raised to memapsin 2 and an alkaline phosphatase conjugated
CC   secondary antibody. The inhibitors may be used in the manufacture of a
CC   medicament for the treatment of Alzheimer's disease since memapsin 2 may
CC   be involved in the cleavage of amyloid precursor protein (APP), and for
CC   determining the substrate side-chain preference in memapsin 2 sub-sites.
CC   The present sequence represents human memapsin 2 (either prepromemapsin 2
CC   or mature memapsin)
XX
SQ   Sequence 501 AA;
XX
Query Match          99.1%; Score 2267; DB 5; Length 501;
Best Local Similarity 99.1%; Pred. No. 9.7e-214;
Matches 429; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
XX
QY   1 TONGIRLPLRSGIGAPLGLINLETDEPEPEPRGSGFVEMVNLKGSGGGYVYMTV 60
DB   22 TONGIRLPLRSGIGAPLGLINLETDEPEPEPRGSGFVEMVNLKGSGGGYVYMTV 81
QY   61 GSPPTLNIIVDTGSSNFAVGAAPHPFLHRYQROLSTYRDLRKGVVPTYTQKMEGEL 120

```

```

DB   82 GSPPTLNIIVDTGSSNFAVGAAPHPFLHRYQROLSTYRDLRKGVVPTYTQKMEGEL 141
QY   121 GTDLVSIPIPGVAVTVRANIAATTESDKPFINGSNMWEGILGLAVAIARPDDSLPFPDSL 180
DB   142 GTDLVSIPIPGVAVTVRANIAATTESDKPFINGSNMWEGILGLAVAIARPDDSLPFPDSL 201
QY   181 VKQTHVPMNLFSLQCGAPPLNQSSEVLASVGGSMIGIDHSLYTGSIMWTPPIREMYE 240
DB   202 VKQTHVPMNLFSLQCGAPPLNQSSEVLASVGGSMIGIDHSLYTGSIMWTPPIREMYE 261
QY   241 VIVVEINQDILKMDCKEYNDKSIYDSGTTNRLPKYFEAAVKSIIKAASSTKEFPDG 300
DB   262 VIVVEINQDILKMDCKEYNDKSIYDSGTTNRLPKYFEAAVKSIIKAASSTKEFPDG 321
QY   301 FWLGEOLVQWQAGTTPMNIIPVISLYLMGEVTNOSFRITLLPOQYLRPEVDVATSDDCY 360
DB   322 FWLGEOLVQWQAGTTPMNIIPVISLYLMGEVTNOSFRITLLPOQYLRPEVDVATSDDCY 381
QY   361 KFAISQSSTGTVMGAVIMEGFVVPDRARRKIGFAVAGCHVDFPTAAVEGFPVTLDME 420
DB   382 KFAISQSSTGTVMGAVIMEGFVVPDRARRKIGFAVAGCHVDFPTAAVEGFPVTLDME 441
QY   421 DCGYNIPTDTEST 433
DB   442 DCGYNIPTDTEST 454
XX
RESULT 15
ABR61928
ID   ABR61928 standard; protein; 501 AA.
XX
AC   ABR61928;
XX
DT   12-SEP-2003 (first entry)
XX
DE   Human memapsin 2 protein.
XX
KW   Memapsin 1; nootropic; neuroprotective; memapsin 2; beta secretase;
KW   beta-amyloid protein; Alzheimer's disease; human; enzyme.
XX
OS   Homo sapiens.
XX
FH   Key Location/Qualifiers
FT   Peptide /note= "signal peptide"
FT   Protein /note= "mature protein"
FT   Peptide /note= "propeptide"
FT   Domain /note= "transmembrane domain"
XX
PN   WO2003039454-A2.
XX
PD   15-MAY-2003.
XX
PF   23-OCT-2002; 2002WO-US034324.
XX
PR   23-OCT-2001; 2001US-0335952P.
PR   27-NOV-2001; 2001US-033545P.
PR   14-JAN-2002; 2002US-0348464P.
PR   14-JAN-2002; 2002US-0348615P.
PR   20-JUN-2002; 2002US-0390804P.
PR   19-JUL-2002; 2002US-0397557P.
PR   19-JUL-2002; 2002US-0397619P.
XX
PA   (OKLA-) OKLAHOMA MEDICAL RES FOUND.
XX
PI   (UNIT ) UNIV ILLINOIS FOUND.
XX
PT   Ghosh AK, Tang J, Bilcer G, Chang W, Hong L, Koelech G, Loy J;
PI   Turner RT;
XX
DR   WPI; 2003-541410/51.

```

DR N-PSDB; ACC84849.

XX New peptide compounds are memapsin beta secretase inhibitors used for  
PT treating Alzheimer's disease.

XX Claim 98; Fig 9; 407pp; English.

CC The invention relates to peptide compounds of specified formula. The  
CC compounds exhibit memapsin 2-beta secretase inhibitory activity relative  
CC to memapsin 1-beta secretase and reduce the accumulation of beta-amyloid  
CC protein. The compounds can be used for treating Alzheimer's disease. The  
CC present sequence represents a human memapsin 2 protein (GenBank Index  
CC (G1):6912266)

XX  
SQ Sequence 501 AA;

Query Match 99.1%; Score 2267; DB 6; Length 501;  
Best Local Similarity 99.1%; Pred. No. 9.7e-214;  
Matches 429; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY	1	TOHGIRLP	LRSGI	GAFLGL	INLETDE	PEEPGR	GRSGF	VEWV	DNLR	KGSG	GGY	VE	MTV	60
DB	22	TOHGIRLP	LRSGI	GAFLGL	INLETDE	PEEPGR	GRSGF	VEWV	DNLR	KGSG	GGY	VE	MTV	81
QY	61	GSP	PQT	LN	IL	VD	TG	SS	N	FA	VG	AA	PH	FL
DB	82	GSP	PQT	LN	IL	VD	TG	SS	N	FA	VG	AA	PH	FL
QY	121	GTD	LV	S	I	P	H	G	V	N	T	A	R	A
DB	142	GTD	LV	S	I	P	H	G	V	N	T	A	R	A
QY	181	V	K	O	T	H	V	P	N	L	F	S	I	O
DB	202	V	K	O	T	H	V	P	N	L	F	S	I	O
QY	241	V	I	I	R	V	E	I	N	G	O	D	L	K
DB	262	V	I	I	R	V	E	I	N	G	O	D	L	K
QY	301	F	M	L	G	E	O	L	V	C	M	O	A	G
DB	322	F	M	L	G	E	O	L	V	C	M	O	A	G
QY	361	K	P	A	I	S	S	T	G	T	W	G	A	V
DB	382	K	P	A	I	S	S	T	G	T	W	G	A	V
QY	421	D	C	G	Y	N	I	P	O	T	D	E	S	T
DB	442	D	C	G	Y	N	I	P	O	T	D	E	S	T

Search completed: August 22, 2005, 00:21:39  
Job time : 131 secs

***This Page Blank (uspto)***





QY 181 VKQTHVNLFSLOLCGAGFPLNOSVULASVGSMTIIGIDHSLYTGSLLMTPIRREMYE 240  
DB 181 VKQTHVNLFSLOLCGAGFPLNOSVULASVGSMTIIGIDHSLYTGSLLMTPIRREMYE 240  
QY 241 VIVRVEINGODLKMDCKEVYNDKSIYDSGTTNLRPLPKVFEAAVKSIIKAASSTKEPPDG 300  
DB 241 VIVRVEINGODLKMDCKEVYNDKSIYDSGTTNLRPLPKVFEAAVKSIIKAASSTKEPPDG 300  
QY 301 FWLGEOLVCWQAGTTPNNIFPVISLYLMGEVTNOSFRITLLPOQYLRPVEDVATSODDCY 360  
DB 301 FWLGEOLVCWQAGTTPNNIFPVISLYLMGEVTNOSFRITLLPOQYLRPVEDVATSODDCY 360  
QY 361 KFAISOSSTGTVMGAVIMEGFYVFDPRARRIGFAVSACHVHDEFRTAAVEGPFVTL DME 420  
DB 361 KFAISOSSTGTVMGAVIMEGFYVFDPRARRIGFAVSACHVHDEFRTAAVEGPFVTL DME 420  
QY 421 DCGYNIPQTDSEST 433  
DB 421 DCGYNIPQTDSEST 433

## RESULT 2

US-09-471-669A-66  
Sequence 66, Application US/09471669A  
Patent No. 6830918  
GENERAL INFORMATION:  
APPLICANT: Anderson, John P.  
APPLICANT: Basl, Guribaj  
APPLICANT: Doane, Minh Tam  
APPLICANT: Frigon, No. 6830918mand  
APPLICANT: John, Varghese  
APPLICANT: Power, Michael  
APPLICANT: Sinha, Sukanto  
APPLICANT: Tatsuno, Gwen  
APPLICANT: Tung, Jay  
APPLICANT: Wang, Shuwen  
APPLICANT: McConlogue, Lisa  
TITLE OF INVENTION: BETA-SECRETASE ENZYME COMPOSITIONS AND METHODS  
FILE REFERENCE: 015270-006430US  
CURRENT APPLICATION NUMBER: US/09/471,669A  
PRIOR FILING DATE: 1999-12-24  
PRIOR APPLICATION NUMBER: US 60/114,408  
PRIOR FILING DATE: 1998-12-31  
PRIOR APPLICATION NUMBER: US 60/119,571  
PRIOR FILING DATE: 1999-02-10  
PRIOR APPLICATION NUMBER: US 60/139,172  
PRIOR FILING DATE: 1999-06-15  
NUMBER OF SEQ ID NOS: 108  
SOFTWARE: Patent Ver. 2.1  
SEQ ID NO 66  
LENGTH: 480  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-471-669A-66

Query Match 99.1%; Score 2267; DB 4; Length 480;  
Best Local Similarity 99.1%; Pred. No. 1.2e-235;

Matches 429; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 TQHGIRLPRLSGGLGAPLGLINLITDEEPPEPRGRSGFVEMVDNLKSGSGGYVEMTV 60  
DB 1 TQHGIRLPRLSGGLGAPLGLIRLPRETDEEPPEPRGRSGFVEMVDNLKSGSGGYVEMTV 60  
QY 61 GSPQQTINLIVDTSSSNFAVGAAPFLHRYRQSLSTYRDLRKGYVYVYTGKMGCEL 120  
DB 61 GSPQQTINLIVDTSSSNFAVGAAPFLHRYRQSLSTYRDLRKGYVYVYTGKMGCEL 120  
QY 121 GTDLVSIIPHGPNVTVRANIAITESDKFPIGNSNMEGIIGLAYAEIARPDLSLEPPDSL 180  
DB 121 GTDLVSIIPHGPNVTVRANIAITESDKFPIGNSNMEGIIGLAYAEIARPDLSLEPPDSL 180

QY 181 VKQTHVNLFSLOLCGAGFPLNOSVULASVGSMTIIGIDHSLYTGSLLMTPIRREMYE 240  
DB 181 VKQTHVNLFSLOLCGAGFPLNOSVULASVGSMTIIGIDHSLYTGSLLMTPIRREMYE 240  
QY 241 VIVRVEINGODLKMDCKEVYNDKSIYDSGTTNLRPLPKVFEAAVKSIIKAASSTKEPPDG 300  
DB 241 VIVRVEINGODLKMDCKEVYNDKSIYDSGTTNLRPLPKVFEAAVKSIIKAASSTKEPPDG 300  
QY 301 FWLGEOLVCWQAGTTPNNIFPVISLYLMGEVTNOSFRITLLPOQYLRPVEDVATSODDCY 360  
DB 301 FWLGEOLVCWQAGTTPNNIFPVISLYLMGEVTNOSFRITLLPOQYLRPVEDVATSODDCY 360  
QY 361 KFAISOSSTGTVMGAVIMEGFYVFDPRARRIGFAVSACHVHDEFRTAAVEGPFVTL DME 420  
DB 361 KFAISOSSTGTVMGAVIMEGFYVFDPRARRIGFAVSACHVHDEFRTAAVEGPFVTL DME 420  
QY 421 DCGYNIPQTDSEST 433  
DB 421 DCGYNIPQTDSEST 433

## RESULT 3

US-09-604-608-2  
Sequence 2, Application US/09604608  
Patent No. 6545127  
GENERAL INFORMATION:  
APPLICANT: Tang, Jordan J.N.  
APPLICANT: Lin, Xinh  
TITLE OF INVENTION: Catalytically Active Recombinant Memapsin and Methods  
FILE REFERENCE: OMR 179  
CURRENT APPLICATION NUMBER: US/09/604,608  
PRIOR FILING DATE: 2000-06-27  
PRIOR APPLICATION NUMBER: 60/141,363  
PRIOR FILING DATE: 1999-06-28  
PRIOR APPLICATION NUMBER: 60/168,060  
PRIOR FILING DATE: 1999-11-30  
PRIOR APPLICATION NUMBER: 60/177,836  
PRIOR FILING DATE: 2000-01-25  
PRIOR APPLICATION NUMBER: 60/178,368  
PRIOR FILING DATE: 2000-01-27  
PRIOR APPLICATION NUMBER: 60/210,292  
PRIOR FILING DATE: 2000-06-08  
NUMBER OF SEQ ID NOS: 31  
SOFTWARE: Patent Ver. 2.1  
SEQ ID NO 2  
LENGTH: 488  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:

OTHER INFORMATION: Purified Memapsin 2  
OTHER INFORMATION: Amino Acids 28-48 are remnant putative propeptide  
OTHER INFORMATION: residues  
OTHER INFORMATION: Amino Acids 58-61, 78, 80, 82-83, 116, 118-121,  
OTHER INFORMATION: 156, 166, 174, 246, 274, 276, 278-281, 283, and  
OTHER INFORMATION: 376-377 are residues in contact with the OM99-2  
OTHER INFORMATION: inhibitor  
OTHER INFORMATION: Amino acids 54-57, 61-68, 73-80, 86-89, 109-111,  
OTHER INFORMATION: 113-118, 123-134, 143-154, 165-168, 198-202, and  
OTHER INFORMATION: 220-224 are N-lobe Beta Strands  
OTHER INFORMATION: Amino Acids 184-191 and 210-217 are N-lobe Helices  
OTHER INFORMATION: Amino acids 237-240, 247-249, 251-256, 259-260,  
OTHER INFORMATION: 273-275, 282-285, 316-318, 331-336, 342-348,  
OTHER INFORMATION: 354-357, 366-370, 372-375, 380-383, 390-395,  
OTHER INFORMATION: 400-405, and 418-420 are C-lobe Beta Strands  
OTHER INFORMATION: Amino Acids 286-299, 307-310, 350-353, 384-387,  
OTHER INFORMATION: and 427-431 are C-lobe Helices  
US-09-604-608-2

Query Match 99.1%; Score 2267; DB 4; Length 488;  
Best Local Similarity 99.1%; Pred. No. 1.2e-235;

Matches 429; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```
QY 1 T0HGIRLPRLRSGLGAPLGLKINLETDEBEPERGRSGFVEMVDNLRGKSGGQYVEMTV 60
DB 9 T0HGIRLPRLRSGLGAPLGLRLPRETDEBEPERGRSGFVEMVDNLRGKSGGQYVEMTV 68
QY 61 GSPPTQNTLIVDTGSSNFVGAAPHPFLHRYOROLSTYRDLRKGVVYPTQGMKEGL 120
DB 69 GSPPTQNTLIVDTGSSNFVGAAPHPFLHRYOROLSTYRDLRKGVVYPTQGMKEGL 128
QY 121 GTDLVSIHPGPNVTVRANIAAITESDKFFINGSNMEGILGLAVAEIARPDLSLEPFDSL 180
DB 129 GTDLVSIHPGPNVTVRANIAAITESDKFFINGSNMEGILGLAVAEIARPDLSLEPFDSL 188
QY 181 VKQTHVFNLFSLQLCGAGFPLNQSEVLASVGSMTIGGIDHSLYTGSLWYTPIRREWYE 240
DB 189 VKQTHVFNLFSLQLCGAGFPLNQSEVLASVGSMTIGGIDHSLYTGSLWYTPIRREWYE 248
QY 241 VIIVRVEINGODLKDCKEYNDKSIYDSGTTNLRPKVFEAAVKSIIKAASSTKFPDG 300
DB 249 VIIVRVEINGODLKDCKEYNDKSIYDSGTTNLRPKVFEAAVKSIIKAASSTKFPDG 308
QY 301 FWLGEOLVCMQAGTTPMNIFFVLSLYLMGEVTVNOSFRITILPQOYLAREVEDVATSQDDCY 360
DB 309 FWLGEOLVCMQAGTTPMNIFFVLSLYLMGEVTVNOSFRITILPQOYLAREVEDVATSQDDCY 368
QY 361 KEAISQSSGTVMGAVIMEGFVYVDFRARKRIGFAVSACHVDEFRTAAVEGPFVTLME 420
DB 369 KEAISQSSGTVMGAVIMEGFVYVDFRARKRIGFAVSACHVDEFRTAAVEGPFVTLME 428
QY 421 DCGYNIPQTDST 433
DB 429 DCGYNIPQTDST 441

RESULT 4
US-09-548-372D-4
; Sequence 4, Application US/09548372D
; Patent No. 6420534
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
; FILE REFERENCE: 29915/62801
; CURRENT APPLICATION NUMBER: US/09/548,372D
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-548-372D-4

Query Match 99.1%; Score 2267; DB 4; Length 501;
Best Local Similarity 99.1%; Pred. No. 1,3e-235;
Matches 429; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 121 GTDLVSIHPGPNVTVRANIAAITESDKFFINGSNMEGILGLAVAEIARPDLSLEPFDSL 180
DB 142 GTDLVSIHPGPNVTVRANIAAITESDKFFINGSNMEGILGLAVAEIARPDLSLEPFDSL 201
QY 181 VKQTHVFNLFSLQLCGAGFPLNQSEVLASVGSMTIGGIDHSLYTGSLWYTPIRREWYE 240
DB 202 VKQTHVFNLFSLQLCGAGFPLNQSEVLASVGSMTIGGIDHSLYTGSLWYTPIRREWYE 261
QY 241 VIIVRVEINGODLKDCKEYNDKSIYDSGTTNLRPKVFEAAVKSIIKAASSTKFPDG 300
DB 262 VIIVRVEINGODLKDCKEYNDKSIYDSGTTNLRPKVFEAAVKSIIKAASSTKFPDG 321
QY 301 FWLGEOLVCMQAGTTPMNIFFVLSLYLMGEVTVNOSFRITILPQOYLAREVEDVATSQDDCY 360
DB 322 FWLGEOLVCMQAGTTPMNIFFVLSLYLMGEVTVNOSFRITILPQOYLAREVEDVATSQDDCY 381
QY 361 KEAISQSSGTVMGAVIMEGFVYVDFRARKRIGFAVSACHVDEFRTAAVEGPFVTLME 420
DB 362 KEAISQSSGTVMGAVIMEGFVYVDFRARKRIGFAVSACHVDEFRTAAVEGPFVTLME 441
QY 421 DCGYNIPQTDST 433
DB 442 DCGYNIPQTDST 454

RESULT 5
US-09-548-367D-4
; Sequence 4, Application US/09548367D
; Patent No. 6440698
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
; FILE REFERENCE: 29915/6280H
; CURRENT APPLICATION NUMBER: US/09/548,367D
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-548-367D-4

Query Match 99.1%; Score 2267; DB 4; Length 501;
Best Local Similarity 99.1%; Pred. No. 1,3e-235;
Matches 429; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
```

```
Db 262 VIIVREINQDGLKMDCKEYNDKSI VDSGTTNLRPKVFEAAVKSIKKAASSTKFPDG 321
QY 301 FWLGEQLVCMQAGTTPWNIFFVLSLYLMGEVYNOSFRITLIPQOYLARVEDVATSODDCY 360
Db 322 FWLGEQLVCMQAGTTPWNIFFVLSLYLMGEVYNOSFRITLIPQOYLARVEDVATSODDCY 381
QY 361 KFAISQSSGTGMGAVIMEGFYVFDRAKRIGFAVSAACHVDEFRTAAVEGPFVTL DME 420
Db 382 KFAISQSSGTGMGAVIMEGFYVFDRAKRIGFAVSAACHVDEFRTAAVEGPFVTL DME 441
QY 421 DCGYNIPQTDST 433
Db 442 DCGYNIPQTDST 454

RESULT 6
US-09-551-853D-4
; Sequence 4, Application US/09551853D
; Patent No. 6500667
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
; FILE REFERENCE: 29915/6280L
; CURRENT APPLICATION NUMBER: US/09/551,853D
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-551-853D-4

Query Match 99.1%; Score 2267; DB 4; Length 501;
Best Local Similarity 99.1%; Pred. No. 1.3e-235;
Matches 429; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 TQHGIRLPRLRSGLGAPLGLIETDEPEPEBGRGSEFVEMVDNLRGSGGGYVEMTV 60
Db 22 TQHGIRLPRLRSGLGAPLGLIETDEPEPEBGRGSEFVEMVDNLRGSGGGYVEMTV 81
QY 61 GSPPTLNLIVDTGSSNFVGAAPHPFLHRYROROLSTYRDLRGVVPYTOGKWEGL 120
Db 82 GSPPTLNLIVDTGSSNFVGAAPHPFLHRYROROLSTYRDLRGVVPYTOGKWEGL 141
QY 121 GTDLVSIPIHGPVNTVRANIAAITESDKPFINSNMEGLIGLAVAIAPDDSLFEPDSL 180
Db 142 GTDLVSIPIHGPVNTVRANIAAITESDKPFINSNMEGLIGLAVAIAPDDSLFEPDSL 201
QY 181 VKQTHVPLNLSIQLGAGFPLNQSEVLASVGSMTIGGIDHSLYTGSLMYTPIRREWEY 240
Db 202 VKQTHVPLNLSIQLGAGFPLNQSEVLASVGSMTIGGIDHSLYTGSLMYTPIRREWEY 261
QY 241 VIIVREINQDGLKMDCKEYNDKSI VDSGTTNLRPKVFEAAVKSIKKAASSTKFPDG 300
Db 262 VIIVREINQDGLKMDCKEYNDKSI VDSGTTNLRPKVFEAAVKSIKKAASSTKFPDG 321
QY 301 FWLGEQLVCMQAGTTPWNIFFVLSLYLMGEVYNOSFRITLIPQOYLARVEDVATSODDCY 360
Db 322 FWLGEQLVCMQAGTTPWNIFFVLSLYLMGEVYNOSFRITLIPQOYLARVEDVATSODDCY 381
QY 361 KFAISQSSGTGMGAVIMEGFYVFDRAKRIGFAVSAACHVDEFRTAAVEGPFVTL DME 420
Db 382 KFAISQSSGTGMGAVIMEGFYVFDRAKRIGFAVSAACHVDEFRTAAVEGPFVTL DME 441
```

```
Db 382 KFAISQSSGTGMGAVIMEGFYVFDRAKRIGFAVSAACHVDEFRTAAVEGPFVTL DME 441
QY 421 DCGYNIPQTDST 433
Db 442 DCGYNIPQTDST 454

RESULT 7
US-09-724-566A-2
; Sequence 2, Application US/09724566A
; Patent No. 6627739
; GENERAL INFORMATION:
; APPLICANT: Anderson, John P.
; APPLICANT: Basl, Gurigbal
; APPLICANT: Doane, Minh Tam
; APPLICANT: Erligon, No. 6627739mand
; APPLICANT: John, Varghese
; APPLICANT: Power, Michael
; APPLICANT: Sinha, Sukanto
; APPLICANT: Tatsuno, Gwen
; APPLICANT: Tung, Jay
; APPLICANT: Wang, Shuwen
; APPLICANT: McConlogue, Lisa
; TITLE OF INVENTION: Beta-Secretase Enzyme Compositions and
; FILE REFERENCE: 228-US-NEWC2
; CURRENT APPLICATION NUMBER: US/09/724,566A
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 09/501,708
; PRIOR FILING DATE: 2000-02-10
; PRIOR APPLICATION NUMBER: 60/119,571
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: 60/139,172
; PRIOR FILING DATE: 1999-06-15
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-566A-2

Query Match 99.1%; Score 2267; DB 4; Length 501;
Best Local Similarity 99.1%; Pred. No. 1.3e-235;
Matches 429; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 TQHGIRLPRLRSGLGAPLGLIETDEPEPEBGRGSEFVEMVDNLRGSGGGYVEMTV 60
Db 22 TQHGIRLPRLRSGLGAPLGLIETDEPEPEBGRGSEFVEMVDNLRGSGGGYVEMTV 81
QY 61 GSPPTLNLIVDTGSSNFVGAAPHPFLHRYROROLSTYRDLRGVVPYTOGKWEGL 120
Db 82 GSPPTLNLIVDTGSSNFVGAAPHPFLHRYROROLSTYRDLRGVVPYTOGKWEGL 141
QY 121 GTDLVSIPIHGPVNTVRANIAAITESDKPFINSNMEGLIGLAVAIAPDDSLFEPDSL 180
Db 142 GTDLVSIPIHGPVNTVRANIAAITESDKPFINSNMEGLIGLAVAIAPDDSLFEPDSL 201
QY 181 VKQTHVPLNLSIQLGAGFPLNQSEVLASVGSMTIGGIDHSLYTGSLMYTPIRREWEY 240
Db 202 VKQTHVPLNLSIQLGAGFPLNQSEVLASVGSMTIGGIDHSLYTGSLMYTPIRREWEY 261
QY 241 VIIVREINQDGLKMDCKEYNDKSI VDSGTTNLRPKVFEAAVKSIKKAASSTKFPDG 300
Db 262 VIIVREINQDGLKMDCKEYNDKSI VDSGTTNLRPKVFEAAVKSIKKAASSTKFPDG 321
QY 301 FWLGEQLVCMQAGTTPWNIFFVLSLYLMGEVYNOSFRITLIPQOYLARVEDVATSODDCY 360
Db 322 FWLGEQLVCMQAGTTPWNIFFVLSLYLMGEVYNOSFRITLIPQOYLARVEDVATSODDCY 381
QY 361 KFAISQSSGTGMGAVIMEGFYVFDRAKRIGFAVSAACHVDEFRTAAVEGPFVTL DME 420
Db 382 KFAISQSSGTGMGAVIMEGFYVFDRAKRIGFAVSAACHVDEFRTAAVEGPFVTL DME 441
```

QY 421 DCGYNIPQTDST 433  
 Db 442 DCGYNIPQTDST 454

## RESULT 8

US-09-416-901B-4  
 ; Sequence 4, Application US/09416901B  
 ; Patent No. 6699671  
 ; GENERAL INFORMATION:  
 ; APPLICANT: GURNEY ET AL.  
 ; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES  
 ; FILE REFERENCE: 29915/6280A  
 ; CURRENT APPLICATION NUMBER: US/09/416,901B  
 ; PRIOR FILING DATE: 1999-10-13  
 ; PRIOR APPLICATION NUMBER: US 60/155,493  
 ; PRIOR FILING DATE: 1999-09-23  
 ; PRIOR APPLICATION NUMBER: US 09/404,133  
 ; PRIOR FILING DATE: 1999-09-23  
 ; PRIOR APPLICATION NUMBER: PCT/US99/20881  
 ; PRIOR FILING DATE: 1999-09-23  
 ; PRIOR APPLICATION NUMBER: US 60/101,594  
 ; PRIOR FILING DATE: 1998-09-24  
 ; NUMBER OF SEQ ID NOS: 72  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 4  
 ; LENGTH: 501  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-09-416-901B-4

Query Match 99.1%; Score 2267; DB 4; Length 501;  
 Best Local Similarity 99.1%; Pred. No. 1,3e-235;

Matches 429; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 TOHGIRLPFRSGAGAPLGLIENLETDSEBEPGRGSGFVEMVDNLRGSGGGYVEMTV 60  
 Db 22 TOHGIRLPFRSGAGAPLGLIENLETDSEBEPGRGSGFVEMVDNLRGSGGGYVEMTV 81  
 QY 61 GSPPTNLIVDTGSSNFAVGAAPHPFLHRYOROLSTYRDLRGVYVPTQKMEGL 120  
 Db 82 GSPPTNLIVDTGSSNFAVGAAPHPFLHRYOROLSTYRDLRGVYVPTQKMEGL 141  
 QY 121 GTDVLSPHGPVNTVRANIAATSDKPFINGSNMEGILGLAYAEIARPDSDLPPFDSL 180  
 Db 142 GTDVLSPHGPVNTVRANIAATSDKPFINGSNMEGILGLAYAEIARPDSDLPPFDSL 201  
 QY 181 VKQTHVPLFSIQLCGAGPFLNQSSEVLASVSGSMIIGIDHSLYTGSMTYTPIRREMYE 240  
 Db 202 VKQTHVPLFSIQLCGAGPFLNQSSEVLASVSGSMIIGIDHSLYTGSMTYTPIRREMYE 261  
 QY 241 VIVRVEINGODLKMDCKEYNDKSIYDSGTNNLRPKVFEAAVKSIRKASTTEKPDG 300  
 Db 262 VIVRVEINGODLKMDCKEYNDKSIYDSGTNNLRPKVFEAAVKSIRKASTTEKPDG 321  
 QY 301 FWLGEOLVCMQAGTTPMNIIPVIVSLYLMGEVYNOSFRITILLPOOYLARPEVDVATSODDCY 360  
 Db 322 FWLGEOLVCMQAGTTPMNIIPVIVSLYLMGEVYNOSFRITILLPOOYLARPEVDVATSODDCY 381  
 QY 361 KFAISQSGTGMGAVIMGEFYVFPDRARRKIGFAVSAHVHDEFRTAAVEGFPVTLDM 420  
 Db 382 KFAISQSGTGMGAVIMGEFYVFPDRARRKIGFAVSAHVHDEFRTAAVEGFPVTLDM 441  
 QY 421 DCGYNIPQTDST 433  
 Db 442 DCGYNIPQTDST 454

RESULT 9  
 US-09-548-376D-4  
 ; Sequence 4, Application US/09548376D

Patent No. 6706485  
 ; GENERAL INFORMATION:  
 ; APPLICANT: GURNEY ET AL.  
 ; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR  
 ; FILE REFERENCE: 29915/6280A  
 ; CURRENT APPLICATION NUMBER: US/09/548,376D  
 ; PRIOR FILING DATE: 2000-04-12  
 ; PRIOR APPLICATION NUMBER: US 60/155,493  
 ; PRIOR FILING DATE: 1999-09-23  
 ; PRIOR APPLICATION NUMBER: US 09/404,133  
 ; PRIOR FILING DATE: 1999-09-23  
 ; PRIOR APPLICATION NUMBER: PCT/US99/20881  
 ; PRIOR FILING DATE: 1999-09-23  
 ; PRIOR APPLICATION NUMBER: US 60/101,594  
 ; PRIOR FILING DATE: 1998-09-24  
 ; NUMBER OF SEQ ID NOS: 73  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 4  
 ; LENGTH: 501  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-09-548-376D-4

Query Match 99.1%; Score 2267; DB 4; Length 501;  
 Best Local Similarity 99.1%; Pred. No. 1,3e-235;

Matches 429; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 TOHGIRLPFRSGAGAPLGLIENLETDSEBEPGRGSGFVEMVDNLRGSGGGYVEMTV 60  
 Db 22 TOHGIRLPFRSGAGAPLGLIENLETDSEBEPGRGSGFVEMVDNLRGSGGGYVEMTV 81  
 QY 61 GSPPTNLIVDTGSSNFAVGAAPHPFLHRYOROLSTYRDLRGVYVPTQKMEGL 120  
 Db 82 GSPPTNLIVDTGSSNFAVGAAPHPFLHRYOROLSTYRDLRGVYVPTQKMEGL 141  
 QY 121 GTDVLSPHGPVNTVRANIAATSDKPFINGSNMEGILGLAYAEIARPDSDLPPFDSL 180  
 Db 142 GTDVLSPHGPVNTVRANIAATSDKPFINGSNMEGILGLAYAEIARPDSDLPPFDSL 201  
 QY 181 VKQTHVPLFSIQLCGAGPFLNQSSEVLASVSGSMIIGIDHSLYTGSMTYTPIRREMYE 240  
 Db 202 VKQTHVPLFSIQLCGAGPFLNQSSEVLASVSGSMIIGIDHSLYTGSMTYTPIRREMYE 261  
 QY 241 VIVRVEINGODLKMDCKEYNDKSIYDSGTNNLRPKVFEAAVKSIRKASTTEKPDG 300  
 Db 262 VIVRVEINGODLKMDCKEYNDKSIYDSGTNNLRPKVFEAAVKSIRKASTTEKPDG 321  
 QY 301 FWLGEOLVCMQAGTTPMNIIPVIVSLYLMGEVYNOSFRITILLPOOYLARPEVDVATSODDCY 360  
 Db 322 FWLGEOLVCMQAGTTPMNIIPVIVSLYLMGEVYNOSFRITILLPOOYLARPEVDVATSODDCY 381  
 QY 361 KFAISQSGTGMGAVIMGEFYVFPDRARRKIGFAVSAHVHDEFRTAAVEGFPVTLDM 420  
 Db 382 KFAISQSGTGMGAVIMGEFYVFPDRARRKIGFAVSAHVHDEFRTAAVEGFPVTLDM 441  
 QY 421 DCGYNIPQTDST 433  
 Db 442 DCGYNIPQTDST 454

## RESULT 10

US-09-794-927A-4  
 ; Sequence 4, Application US/09794927A  
 ; Patent No. 6727074

GENERAL INFORMATION:  
 ; APPLICANT: Gurney et al.  
 ; TITLE OF INVENTION: Alzheimer's Disease Secretase, APP Substrates Therefor and Uses  
 ; FILE REFERENCE: 29915/6280FG  
 ; CURRENT APPLICATION NUMBER: US/09/794,927A  
 ; CURRENT FILING DATE: 2001-02-27

```
/ PRIOR APPLICATION NUMBER: 09/416,901
/ PRIOR FILING DATE: 1999-10-13
/ PRIOR APPLICATION NUMBER: 60/155,493
/ PRIOR FILING DATE: 1999-09-23
/ PRIOR APPLICATION NUMBER: 09/404,133
/ PRIOR FILING DATE: 1999-09-23
/ PRIOR APPLICATION NUMBER: PCT/US99/20881
/ PRIOR FILING DATE: 1999-09-23
/ PRIOR APPLICATION NUMBER: 60/101,594
/ PRIOR FILING DATE: 1998-09-24
/ NUMBER OF SEQ ID NOS: 74
/ SOFTWARE: Patent In Ver. 2.0
/ SEQ ID NO 4
/ LENGTH: 501
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-794-927a-4
```

```
Query Match          99.1%; Score 2267; DB 4; Length 501;
Best Local Similarity 99.1%; Pred. No. 1.3e-235;
Matches 429; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 TQHGIRLPRLRSGLGAPLGLINLETDEEPEEPRGRSGFVEMVDNLKSGGQGYVEMTV 60
DB 22 TQHGIRLPRLRSGLGAPLGLRLPRETDEEPEEPRGRSGFVEMVDNLKSGGQGYVEMTV 81
QY 61 GSPQQTNLIVDTSSNFAVGAAPHPFLHRYRQSLSTYRDLRKGVYVPTQKMEGEL 120
DB 82 GSPQQTNLIVDTSSNFAVGAAPHPFLHRYRQSLSTYRDLRKGVYVPTQKMEGEL 141
QY 121 GTDLVSIIPHGNVTVRANIAITSDKFFINGSNMEGILGLAYAEIARPDLSLEPPFDSL 180
DB 142 GTDLVSIIPHGNVTVRANIAITSDKFFINGSNMEGILGLAYAEIARPDLSLEPPFDSL 201
QY 181 VKQTHVPLBSLQCGAFPLNOSSEVLASVGSMTIGIDHSLYTGSLWYTPIRREWYVE 240
DB 202 VKQTHVPLBSLQCGAFPLNOSSEVLASVGSMTIGIDHSLYTGSLWYTPIRREWYVE 261
QY 241 VIIVRVEINGQDLKMDCKEYNDKSIYDSGTTNLRPKVFEAAVKSIRKASSTKFPDG 300
DB 262 VIIVRVEINGQDLKMDCKEYNDKSIYDSGTTNLRPKVFEAAVKSIRKASSTKFPDG 321
QY 301 FWLGEQLVCMQAGTTTPWNIIPVVISLYLMEVTVNQSFRTITLPOQYLRPEVDVATSDDDCY 360
DB 322 FWLGEQLVCMQAGTTTPWNIIPVVISLYLMEVTVNQSFRTITLPOQYLRPEVDVATSDDDCY 381
QY 361 KFAISQSTGTVMGAVIMEGFYVFPDARKRIGFAVSACHVDEFTAAVEGFPVTLTME 420
DB 382 KFAISQSTGTVMGAVIMEGFYVFPDARKRIGFAVSACHVDEFTAAVEGFPVTLTME 441
QY 421 DCGYNIPQTDST 433
DB 442 DCGYNIPQTDST 454
```

```
RESULT 11
US-09-548-373D-4
/ Sequence 4, Application US/09548373D
/ Patent No. 6737510
/ GENERAL INFORMATION:
/ APPLICANT: GURNEY ET AL.
/ TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
/ TITLE OF INVENTION: THEROPF
/ FILE REFERENCE: 29915/62808
/ CURRENT APPLICATION NUMBER: US/09/548,373D
/ CURRENT FILING DATE: 2000-04-12
/ PRIOR APPLICATION NUMBER: US 60/155,493
/ PRIOR FILING DATE: 1999-09-23
/ PRIOR APPLICATION NUMBER: US 09/404,133
/ PRIOR FILING DATE: 1999-09-23
/ PRIOR APPLICATION NUMBER: PCT/US99/20881
/ PRIOR FILING DATE: 1999-09-23
/ PRIOR APPLICATION NUMBER: US 60/101,594
```

```
/ PRIOR FILING DATE: 1998-09-24
/ NUMBER OF SEQ ID NOS: 73
/ SOFTWARE: Patent In version 3.1
/ SEQ ID NO 4
/ LENGTH: 501
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-548-373D-4
```

```
Query Match          99.1%; Score 2267; DB 4; Length 501;
Best Local Similarity 99.1%; Pred. No. 1.3e-235;
Matches 429; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 TQHGIRLPRLRSGLGAPLGLINLETDEEPEEPRGRSGFVEMVDNLKSGGQGYVEMTV 60
DB 22 TQHGIRLPRLRSGLGAPLGLRLPRETDEEPEEPRGRSGFVEMVDNLKSGGQGYVEMTV 81
QY 61 GSPQQTNLIVDTSSNFAVGAAPHPFLHRYRQSLSTYRDLRKGVYVPTQKMEGEL 120
DB 82 GSPQQTNLIVDTSSNFAVGAAPHPFLHRYRQSLSTYRDLRKGVYVPTQKMEGEL 141
QY 121 GTDLVSIIPHGNVTVRANIAITSDKFFINGSNMEGILGLAYAEIARPDLSLEPPFDSL 180
DB 142 GTDLVSIIPHGNVTVRANIAITSDKFFINGSNMEGILGLAYAEIARPDLSLEPPFDSL 201
QY 181 VKQTHVPLBSLQCGAFPLNOSSEVLASVGSMTIGIDHSLYTGSLWYTPIRREWYVE 240
DB 202 VKQTHVPLBSLQCGAFPLNOSSEVLASVGSMTIGIDHSLYTGSLWYTPIRREWYVE 261
QY 241 VIIVRVEINGQDLKMDCKEYNDKSIYDSGTTNLRPKVFEAAVKSIRKASSTKFPDG 300
DB 262 VIIVRVEINGQDLKMDCKEYNDKSIYDSGTTNLRPKVFEAAVKSIRKASSTKFPDG 321
QY 301 FWLGEQLVCMQAGTTTPWNIIPVVISLYLMEVTVNQSFRTITLPOQYLRPEVDVATSDDDCY 360
DB 322 FWLGEQLVCMQAGTTTPWNIIPVVISLYLMEVTVNQSFRTITLPOQYLRPEVDVATSDDDCY 381
QY 361 KFAISQSTGTVMGAVIMEGFYVFPDARKRIGFAVSACHVDEFTAAVEGFPVTLTME 420
DB 382 KFAISQSTGTVMGAVIMEGFYVFPDARKRIGFAVSACHVDEFTAAVEGFPVTLTME 441
QY 421 DCGYNIPQTDST 433
DB 442 DCGYNIPQTDST 454
```

```
RESULT 12
US-09-795-847B-4
/ Sequence 4, Application US/09795847B
/ Patent No. 6753163
/ GENERAL INFORMATION:
/ APPLICANT: Gurney, Mark E.
/ APPLICANT: Bienkowski, Michael J.
/ APPLICANT: Heintzkeon, Robert L.
/ APPLICANT: Patodi, Luis A.
/ APPLICANT: Yan, Riqiang
/ TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND USES
/ TITLE OF INVENTION: THEROPF
/ FILE REFERENCE: 28341/62808
/ CURRENT APPLICATION NUMBER: US/09/795,847B
/ CURRENT FILING DATE: 2001-02-28
/ PRIOR APPLICATION NUMBER: 09/416,901
/ PRIOR FILING DATE: 1999-10-13
/ PRIOR APPLICATION NUMBER: 60/155,493
/ PRIOR FILING DATE: 1999-09-23
/ PRIOR APPLICATION NUMBER: 09/404,133
/ PRIOR FILING DATE: 1999-09-23
/ PRIOR APPLICATION NUMBER: PCT/US99/20881
/ PRIOR FILING DATE: 1999-09-23
/ PRIOR APPLICATION NUMBER: 60/101,594
/ PRIOR FILING DATE: 1998-09-24
/ NUMBER OF SEQ ID NOS: 74
/ SOFTWARE: Patent In Ver. 2.0
```

SEQ ID NO 4  
 LENGTH: 501  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-09-795-847B-4

Query Match 99.1%; Score 2267; DB 4; Length 501;  
 Best Local Similarity 99.1%; Pred. No. 1.3e-235;  
 Matches 429; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

1 TGHGRLPLRSGLGAPLGLRLINLETDEBEPGRGRGSFVEMVDNLKSGSGGYVEMTV 60  
 22 TGHGRLPLRSGLGAPLGLRLPRETDEBEPGRGRGSFVEMVDNLKSGSGGYVEMTV 81  
 61 GSPPTLNLIVDTGSSNFVGAAPHPFLHRYQROLSTYRDLKGVVYPYTOGKMEGL 120  
 82 GSPPTLNLIVDTGSSNFVGAAPHPFLHRYQROLSTYRDLKGVVYPYTOGKMEGL 141  
 121 GTDLVSIPIHGNVTVRANIAAITSBDFPINGSNMEGILGLAYAEIARPDLSLEFPDLS 180  
 142 GTDLVSIPIHGNVTVRANIAAITSBDFPINGSNMEGILGLAYAEIARPDLSLEFPDLS 201  
 181 VKQTHVPMFLSLQCGAFPLNOSVVLASVGSMTIGSIDSLYTGSLMTYPIRREMYE 240  
 202 VKQTHVPMFLSLQCGAFPLNOSVVLASVGSMTIGSIDSLYTGSLMTYPIRREMYE 261  
 241 VIVVEINGODLKMDCKEYNDKSIYDSGTTNLRPKVFEAAVKSIIKAASSTKFPDG 300  
 262 VIVVEINGODLKMDCKEYNDKSIYDSGTTNLRPKVFEAAVKSIIKAASSTKFPDG 321  
 301 FWLGEQLVCWQAGTTPMNIIPVLSLYLMEGVYTNOSFRITILPQYLRPEVEDVATSQDDCY 360  
 322 FWLGEQLVCWQAGTTPMNIIPVLSLYLMEGVYTNOSFRITILPQYLRPEVEDVATSQDDCY 381  
 361 KPAISQSTGYTMGAVIMEGFVYVFPDRARKRIGFAVSACHVDEPRTAABGPFVTLDM 420  
 382 KPAISQSTGYTMGAVIMEGFVYVFPDRARKRIGFAVSACHVDEPRTAABGPFVTLDM 441  
 421 DCGYNIPQTDST 433  
 442 DCGYNIPQTDST 454

## RESULT 13

US-09-548-366F-4  
 Sequence 4; Application US/09548366F  
 Patent No. 6797487  
 GENERAL INFORMATION:  
 APPLICANT: GURNEY ET AL.  
 TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES  
 TITLE OF INVENTION: THEREOF  
 FILE REFERENCE: 29915/6280J  
 CURRENT APPLICATION NUMBER: US/09/548,366F  
 CURRENT FILING DATE: 2000-04-12  
 PRIOR APPLICATION NUMBER: US 60/155,493  
 PRIOR FILING DATE: 1999-09-23  
 PRIOR APPLICATION NUMBER: US 09/404,133  
 PRIOR FILING DATE: 1999-09-23  
 PRIOR APPLICATION NUMBER: PCT/US99/20881  
 PRIOR FILING DATE: 1999-09-23  
 PRIOR APPLICATION NUMBER: US 60/101,594  
 PRIOR FILING DATE: 1998-09-24  
 NUMBER OF SEQ ID NOS: 73  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 4  
 LENGTH: 501  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-09-548-366F-4

Query Match 99.1%; Score 2267; DB 4; Length 501;  
 Best Local Similarity 99.1%; Pred. No. 1.3e-235;  
 Matches 429; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

1 TGHGRLPLRSGLGAPLGLRLINLETDEBEPGRGRGSFVEMVDNLKSGSGGYVEMTV 60  
 22 TGHGRLPLRSGLGAPLGLRLPRETDEBEPGRGRGSFVEMVDNLKSGSGGYVEMTV 81  
 61 GSPPTLNLIVDTGSSNFVGAAPHPFLHRYQROLSTYRDLKGVVYPYTOGKMEGL 120  
 82 GSPPTLNLIVDTGSSNFVGAAPHPFLHRYQROLSTYRDLKGVVYPYTOGKMEGL 141  
 121 GTDLVSIPIHGNVTVRANIAAITSBDFPINGSNMEGILGLAYAEIARPDLSLEFPDLS 180  
 142 GTDLVSIPIHGNVTVRANIAAITSBDFPINGSNMEGILGLAYAEIARPDLSLEFPDLS 201  
 181 VKQTHVPMFLSLQCGAFPLNOSVVLASVGSMTIGSIDSLYTGSLMTYPIRREMYE 240  
 202 VKQTHVPMFLSLQCGAFPLNOSVVLASVGSMTIGSIDSLYTGSLMTYPIRREMYE 261  
 241 VIVVEINGODLKMDCKEYNDKSIYDSGTTNLRPKVFEAAVKSIIKAASSTKFPDG 300  
 262 VIVVEINGODLKMDCKEYNDKSIYDSGTTNLRPKVFEAAVKSIIKAASSTKFPDG 321  
 301 FWLGEQLVCWQAGTTPMNIIPVLSLYLMEGVYTNOSFRITILPQYLRPEVEDVATSQDDCY 360  
 322 FWLGEQLVCWQAGTTPMNIIPVLSLYLMEGVYTNOSFRITILPQYLRPEVEDVATSQDDCY 381  
 361 KPAISQSTGYTMGAVIMEGFVYVFPDRARKRIGFAVSACHVDEPRTAABGPFVTLDM 420  
 382 KPAISQSTGYTMGAVIMEGFVYVFPDRARKRIGFAVSACHVDEPRTAABGPFVTLDM 441  
 421 DCGYNIPQTDST 433  
 442 DCGYNIPQTDST 454

## RESULT 14

US-09-949-016-6810  
 Sequence 6810; Application US/09949016  
 Patent No. 6812339  
 GENERAL INFORMATION:  
 APPLICANT: VENTER, J. Craig et al.  
 TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
 TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
 FILE REFERENCE: C100107  
 CURRENT APPLICATION NUMBER: US/09/949,016  
 CURRENT FILING DATE: 2000-04-14  
 PRIOR APPLICATION NUMBER: 60/241,755  
 PRIOR FILING DATE: 2000-10-20  
 PRIOR APPLICATION NUMBER: 60/237,768  
 PRIOR FILING DATE: 2000-10-03  
 PRIOR APPLICATION NUMBER: 60/231,498  
 PRIOR FILING DATE: 2000-09-08  
 NUMBER OF SEQ ID NOS: 207012  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 6810  
 LENGTH: 501  
 TYPE: PRT  
 ORGANISM: Human  
 US-09-949-016-6810

Query Match 99.1%; Score 2267; DB 4; Length 501;  
 Best Local Similarity 99.1%; Pred. No. 1.3e-235;  
 Matches 429; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

1 TGHGRLPLRSGLGAPLGLRLINLETDEBEPGRGRGSFVEMVDNLKSGSGGYVEMTV 60  
 22 TGHGRLPLRSGLGAPLGLRLPRETDEBEPGRGRGSFVEMVDNLKSGSGGYVEMTV 81  
 61 GSPPTLNLIVDTGSSNFVGAAPHPFLHRYQROLSTYRDLKGVVYPYTOGKMEGL 120  
 82 GSPPTLNLIVDTGSSNFVGAAPHPFLHRYQROLSTYRDLKGVVYPYTOGKMEGL 141  
 121 GTDLVSIPIHGNVTVRANIAAITSBDFPINGSNMEGILGLAYAEIARPDLSLEFPDLS 180

```
Db 142 GTDLVSIPIHGPVTVRANIAITBESDKFFINGSNWEGILGLAVAEIARPDLSLEPPFDSL 201
QY 181 VKQTHVNLPSLQICGAGFPLNOSSEVTLASVGSMTIGIDHSLYTSGSLWYTPPIRREWYE 240
Db 202 VKQTHVNLPSLQICGAGFPLNOSSEVTLASVGSMTIGIDHSLYTSGSLWYTPPIRREWYE 261
QY 241 VVIYRVEINGODLKMDCKEVNYDYSIDSGTTNLRPKVPEAAVKSIIKAASSTKPPDG 300
Db 262 VVIYRVEINGODLKMDCKEVNYDYSIDSGTTNLRPKVPEAAVKSIIKAASSTKPPDG 321
QY 301 FWLGEQLVCWQAGTTPMNIFFVISLYLMGEVTNOSFRITILLPOQYLRPVEDVATSODDCY 360
Db 322 FWLGEQLVCWQAGTTPMNIFFVISLYLMGEVTNOSFRITILLPOQYLRPVEDVATSODDCY 381
QY 361 KFAISQSTGTVMGAVIMEGFYVFPDARKRIGFAVSACHVHDEFRTAAVEGPFVTLDME 420
Db 382 KFAISQSTGTVMGAVIMEGFYVFPDARKRIGFAVSACHVHDEFRTAAVEGPFVTLDME 441
QY 421 DCGYNIPTDEST 433
Db 442 DCGYNIPTDEST 454
```

```
RESULT 15
US-09-548-368D-4
; Sequence 4, Application US/09548368D
; Patent No. 6825023
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
; FILE REFERENCE: 29915/6280C
; CURRENT APPLICATION NUMBER: US/09/548,368D
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-548-368D-4
```

```
Query Match 99.1%; Score 2267; DB 4; Length 501;
Best Local Similarity 99.1%; Pred. No. 1,3e-235;
Matches 429; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 TOHGIRLPLRSGLGAPLGLIEINLETDBRPERGRGSPVEMVNDLRKSGGQYIVEMTV 60
Db 22 TOHGIRLPLRSGLGAPLGLRPRETDEBPGRGSGFVEMVNDLRKSGGQYIVEMTV 81
QY 61 GSPPTNLINLVDTGSSNFAVGAAHPFLHRYQROLSTYRDLRKGVVVPYTOGKMEGEL 120
Db 82 GSPPTNLINLVDTGSSNFAVGAAHPFLHRYQROLSTYRDLRKGVVVPYTOGKMEGEL 141
QY 121 GTDLVSIPIHGPVTVRANIAITBESDKFFINGSNWEGILGLAVAEIARPDLSLEPPFDSL 180
Db 142 GTDLVSIPIHGPVTVRANIAITBESDKFFINGSNWEGILGLAVAEIARPDLSLEPPFDSL 201
QY 181 VKQTHVNLPSLQICGAGFPLNOSSEVTLASVGSMTIGIDHSLYTSGSLWYTPPIRREWYE 240
Db 202 VKQTHVNLPSLQICGAGFPLNOSSEVTLASVGSMTIGIDHSLYTSGSLWYTPPIRREWYE 261
QY 241 VVIYRVEINGODLKMDCKEVNYDYSIDSGTTNLRPKVPEAAVKSIIKAASSTKPPDG 300
Db 262 VVIYRVEINGODLKMDCKEVNYDYSIDSGTTNLRPKVPEAAVKSIIKAASSTKPPDG 321
```

```
QY 301 FWLGEQLVCWQAGTTPMNIFFVISLYLMGEVTNOSFRITILLPOQYLRPVEDVATSODDCY 360
Db 322 FWLGEQLVCWQAGTTPMNIFFVISLYLMGEVTNOSFRITILLPOQYLRPVEDVATSODDCY 381
QY 361 KFAISQSTGTVMGAVIMEGFYVFPDARKRIGFAVSACHVHDEFRTAAVEGPFVTLDME 420
Db 382 KFAISQSTGTVMGAVIMEGFYVFPDARKRIGFAVSACHVHDEFRTAAVEGPFVTLDME 441
QY 421 DCGYNIPTDEST 433
Db 442 DCGYNIPTDEST 454
```

```
Search completed: August 22, 2005, 00:26:27
Job time : 67 sec
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 22, 2005, 00:25:14 ; Search time 5781 Seconds  
(without alignment)  
2851.034 Million cell updates/sec

Title: US-10-726-967a-84

Perfect score: 2288

Sequence: 1 TGHGIRLPURSGLGAPGL.....FTLDMEDCGNIPQTDST 433

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3423544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODE=frame+ p2n.model -DEV=xlp  
-Q/cgcn2.1/USPRO.spool.P/US10726967/runat.19082005.201240.29636/app\_query.faeta\_1.583  
-DB=EST -QFMT=fastap -SUFFIX=tbl -MIMATCH=0.1 -LOOPT=0 -LOOPTXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=diolunet2 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pcio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10726967 @CGCN 1.1 5180 @runat.19082005.201240.29636 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -MARN\_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEX=7

Database : EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_est4:\*  
5: gb\_est5:\*  
6: gb\_est6:\*  
7: gb\_est7:\*  
8: gb\_est8:\*  
9: gb\_est9:\*  
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2267	99.1	1506	9 AY417360	Homo sapi
2	2222	97.1	1445	9 AY417361	Pan trogl
3	2216	96.9	1506	9 AY417362	Mus muscu
4	2216	96.9	3634	3 AK041285	Mus muscu
5	2216	96.9	3859	3 AK044464	Mus muscu
6	2216	96.9	3877	3 AK031112	Mus muscu
7	2215	96.8	4046	3 AK049626	Mus muscu
8	2215	96.8	4101	3 AK046175	Mus muscu
9	2189	95.7	3880	3 AK080498	Mus muscu

10	2181	95.3	3805	3 AK082230	Mus muscu
11	2023	88.4	4048	3 AK082317	Mus muscu
12	1537	67.2	3147	3 AK079814	Mus muscu
13	1425	62.3	928	5 B0557282	AGENCOURT
14	1375.5	60.1	1001	5 B0128383	603113984
15	1365	59.7	950	5 B0502697	AGENCOURT
16	1364	59.6	802	5 B0613719	UI-M-EWO-
17	1315.5	57.5	846	7 CK021791	AGENCOURT
18	1273.5	55.7	926	5 BX433493	BX433493
19	1257	54.9	862	4 B1157614	602922587
20	1240.5	54.2	813	7 CN224123	WLA053B12
21	1163	50.8	2504	4 AK014390	Mus muscu
22	1158.5	50.6	720	4 BG288435	602383404
23	1136	49.7	1968	3 CR622762	FW1-1eng
24	1135	49.6	840	2 BB885725	UI-M-FCO-
25	1122	49.0	836	7 C0382905	AGENCOURT
26	1120	49.0	3703	3 AK078770	Mus muscu
27	1115	48.7	1810	3 AK052309	603625683
28	1100.5	48.1	730	4 BM048418	UI-M-GHO-
29	1099	48.0	631	7 CFS39391	UI-M-FCO-
30	1093	47.8	624	5 B0708350	AF150387
31	1048	45.8	1114	1 AF150387	AF150387
32	1036	45.3	684	1 AV725513	AV725513
33	1036	45.3	904	2 BE378929	601237528
34	1007	44.0	808	7 CK833062	4056853 B
35	1005.5	43.9	633	5 B0073919	1n26c07.Y
36	997	43.6	789	2 BF204338	601866411
37	992	43.4	635	2 BE296749	601175134
38	972	42.5	797	5 B0193314	AGENCOURT
39	956.5	41.8	605	2 BE793449	601588510
40	954	41.7	826	5 B0182102	AGENCOURT
41	941	41.1	526	6 BY713531	K0314H01-
42	923	40.3	529	6 CA563516	AV667139
43	912	39.9	942	1 AV667139	601868788
44	907.5	39.7	939	2 BF203806	601868788
45	892	39.0	513	5 B0637035	he04912.Y

## ALIGNMENTS

RESULT 1	AY417360	1506 bp	DNA	linear	GSS 17-DEC-2003
LOCUS	AY417360				
DEFINITION	Homo sapiens BACE gene, VIRTUAL TRANSCRIPT, partial sequence.				
ACCESSION	AY417360				
VERSION	AY417360.1	GI:39773320			
KEYWORDS	GSS.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shtinsky,J.J., Adams,M.D. and Cargill,M.				
TITLE	Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios				
JOURNAL	Science	302 (5652),	1960-1963	(2003)	
PUBMED	14671302				
REFERENCE	2 (bases 1 to 1506)				
AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shtinsky,J.J., Adams,M.D. and Cargill,M.				
TITLE	Direct Substitution				
JOURNAL	Submitted (16-NOV-2003)	Celera Genomics,	45 West Gude Drive,		
COMMENT	Rockville, MD 20850, USA				
FEATURES	This sequence was made by sequencing genomic exons and ordering them based on alignment.				
source	1. 1506				

gene

/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
<1..>1506  
/gene="BACE"  
/locus\_tag="HCM6198"

## ORIGIN

## Alignment Scores:

Pred. No.:	1.2e-245	Length:	1506
Score:	2267.00	Matches:	429
Percent Similarity:	99.31%	Conservative:	1
Best Local Similarity:	99.08%	Mismatches:	3
Query Match:	99.08%	Indels:	0
DB:	9	Gaps:	0

US-10-726-967a-84 (1-433) x AY17360 (1-1506)

```

QY      1  ThrGlnHisGlyIleArgLeuProLeuArgSerGlyLeuGlyValAlaProLeuGlyLeu 20
Db      64  ACCAGCAGCGCATCCGGCTGCCCTGGCAGCGGGCTGGGGGGCCCCCTGGGGCTG 123

QY      21  GluIleAsnLeuGluThrArgGluGluProGluGluProGlyValArgArgGlySerPheVal 40
Db      124  CGGCTGCCCGGAGAGCCGACGAAAGCCCGAGAGAGCCCGCGGAGGGCGAGCTTTGTG 183

QY      41  GluMetValAspAsnLeuArgGlyLysSerGlyGlnGlyTyrTyrValGluMetThrVal 60
Db      184  GAGATGGTGGACAACTCGAGGGGCGAAGTCCGGGCGAGGCTACTACGTGAGATGACCGTG 243

QY      61  GlySerProProGlnThrLeuAsnIleLeuValAspThrGlySerSerAsnPheAlaVal 80
Db      244  GGCAGGCCCGCCGAGAGCGCTCAACATCTGGTGTGATACAGGCGAGCAGTAACTTGGACGTG 303

QY      81  GlyValAlaAspHisProPheLeuHisArgTyrTyrGlnArgGlnLeuSerSerThrTyr 100
Db      304  GGTGTGCCCCCGACCCCTTCTTCGATCGCTACTACAGAGGAGGTGTCCAGCACATAC 363

QY      101  ArgAspLeuArgGlyValTyrValProTyrThrGlnGlyLysTrpGluGlyLeu 120
Db      364  CGGAGCTCCGGAAAGGTGTGTGTGTGCCCTACACCCAGGGGCAAGTGGGAAAGGGAGCTG 423

QY      121  GlyThrAspLeuValSerIleProHisGlyProAsnValThrValArgAlaAsnIleAla 140
Db      424  GGCACCGAGCTGTAGATGCCCCCAATGCCCCCAAGTCACTGTGCTGCCAACATTTGCT 483

QY      141  AlaIleThrGlySerAspLysPhePheIleAsnGlySerAsnTrpGluGlyIleLeuGly 160
Db      484  GCCATCAGTCAATCAGACAAAGTTCTTTCATCAACGGCTCCAACTGGGAAAGGCAATCTG 543

QY      161  LeuAlaTyrAlaGluIleAlaArgProAspAspSerLeuGluProPheAspSerLeu 180
Db      544  CTGGGCTTAAGTGTGAGTTGCCAGGCTCGACCTCCCTGAGACCTTTCTTTCGACTCTG 603

QY      181  ValLysGlnThrHisValProAsnLeuPheSerLeuGlnLeuCysGlyValGlyPhePro 200
Db      604  GTTAAAGCAGACCCAGCTTCCCAACCTCTTCTCCGACGCTTGTGGTGGCTTCCCC 663

QY      201  LeuAsnGlnSerGluValLeuAlaSerValGlyLysMetIleIleGlyIleAsp 220
Db      664  CTCACACCAAGTCTGAGTGTGCTCTGTCCGAGGAGGACATGATCATTTGGAGGTATGAC 723

QY      221  HisSerLeuTyrThrGlySerLeuTrpTyrThrProIleArgArgGluTrpTyrTyrGlu 240
Db      724  CACTGCTGTACACAGGAGATCTCTGTGTATACACCACTCCGCGGAGATGTATTATGAG 783

QY      241  ValIleIleValArgValGluIleAsnGlyGlnAspLeuLysMetAspCysValGluTyr 260
Db      784  GTGATCATTTGCGGGGTGAGATCAATGACAGATCTGAAATAAGGACTGCAAGGAGATC 843

QY      261  AsnTyrAspLysSerIleValAspSerGlyThrThrAsnLeuArgLeuProLysVal 280
Db      844  AACTATGACAAAGACATGTGTGAGAGTGGACACCAACCTTTCGTTTCCCAAGAAAGTG 903

```

```

QY      281  PheGluAlaAlaValLysSerIleLysValAlaAspSerThrGluLysPheProAspGly 300
Db      904  TTTGAAGCTGCAAGTCAATTCATCAAGGACAGCTCTCTCCAGGAAAGTTCCCTGATGCT 963

QY      301  PheTrpLeuGlyGluGlnLeuValCysTrpGlnAlaGlyThrThrProTrpAsnIlePhe 320
Db      964  TTCTGGCTAGAGAGACACTGGTGTGCTGGCAGACAGACACACCCCTTGGAAACATTTTC 1023

QY      321  ProValIleSerLeuTyrLeuMetGlyGluValThrAsnGlnSerPheArgIleThrIle 340
Db      1024  CCAGTCATCTCACTTACTTAAATGGGTAGGTTTCAACACAGTCTCTCCGATACCAATC 1083

QY      341  LeuProGlnGlnTyrLeuArgProValGluAspValAlaThrSerGlnAspAspCysTyr 360
Db      1084  CTTCCGCGCAATACCTGCGCGCCAGTGGAGATGTGGCCACAGTCCCAAGACGACTTTAC 1143

QY      361  LysPheAlaIleSerGlnSerSerThrGlyThrValMetGlyValAlaIleMetGluGly 380
Db      1144  AAGTTTGCATCTCAAGTCAATCCAGCGGACGTGTATGGAGCTGTATCATGAGAGGC 1203

QY      381  PheTyrValValPheAspArgAlaArgLysArgIleGlyPheAlaValSerAlaCysHis 400
Db      1204  TTCTACGTTGTCTTTGATCCGGGCCGAAAGCAATTGGCTTGTGTACAGCGCTTGGCAT 1263

QY      401  ValHisAspGluPheArgThrAlaAlaValGluGlyProPheValThrLeuAspMetGlu 420
Db      1264  GTGCACATATGATTCAGACCGGACGCGGTGAAAGGCCCTTTGTGTCACTTGGACATGGAA 1323

QY      421  AspCysGlyTyrAsnIleProGlnThrAspGlySerThr 433
Db      1324  GACTGTGGCTACAACTTCCACAGACGATGATCAACC 1362

RESULT 2
AY17361      1445 bp      DNA      linear      GSS 17-DEC-2003
LOCUS      Pan troglodytes BACE gene, VIRUAL TRANSCRIPT, partial sequence,
DEFINITION      genomic survey sequence.
ACCESSION      AY17361
VERSION      AY17361.1 GI:39773321
KEYWORDS      GSS.
SOURCE      Pan troglodytes (chimpanzee)
ORGANISM      Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
REFERENCE      Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
AUTHORS      Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
              Ferritera,S., Wang,G., Zheng,X.H., White,T.J., Sniinsky,J.J.,
              Adams,M.D. and Cargill,M.
              Inferring nonneutral evolution from human-chimp-mouse orthologous
              gene trios
              Science 302 (5652), 1960-1963 (2003)
JOURNAL      Science 302 (5652), 1960-1963 (2003)
PUBMED      14671302
REFERENCE      Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
AUTHORS      Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
              Ferritera,S., Wang,G., Zheng,X.H., White,T.J., Sniinsky,J.J.,
              Adams,M.D. and Cargill,M.
              Direct Submision
              Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
              Rockville, MD 20850, USA
              This sequence was made by sequencing genomic exons and ordering
              them based on alignment.
              Location/Qualifiers
FEATURES
     source
         1..1445
         /organism="Pan troglodytes"
         /mol_type="genomic DNA"
         /db_xref="taxon:9598"
         <1..>1445
         /gene="BACE"
         /locus_tag="HCM6198"
ORIGIN

```

## Alignment Scores:

Pred. No.: 1,43e-240 Length: 1445  
 Score: 2222.00 Matches: 422  
 Percent Similarity: 97.69% Conservative: 1  
 Best Local Similarity: 97.46% Mismatches: 10  
 Query Match: 97.12% Indels: 0  
 DB: 9 Gaps: 0

US-10-726-967a-84 (1-433) x AY417361 (1-1445)

QY 1 ThrGlnHsAglyTlLeaArgLeuProLeuAgsSerGlyLeuGlyGlyAlaProLeuGlyLeu 20  
 DB 8 ACCCGACAGCGGATCCGGCTGCCNNNNNNACCGGGCTGGGNNNGCCCCCTGGAGCTG 67  
 QY 21 GtUleamLeuGltThraPgluGluPProGluGluPProGlyAraGrglySerPheVal 40  
 DB 68 CGGCTGCCCCGGAGACCAAGAGGCCGAGAGGCCGGCCGAGAGGGGCGCTTGTG 127  
 QY 41 GluMetValAspAsnLeuArglyLySerGlyGlyGlyTyrTyrValGluMetThrVal 60  
 DB 128 GAGATGTGACAACTGAGGGCAAGTCGGGGCAGAGNCTACTACGTGAGATTNNCGTG 187  
 QY 61 GlySerProProGltThreLeuAnlleLeuValAspThrGlySerSerAsnPheAlaVal 80  
 DB 188 GGCAGCNCNCCCGCAGAGCTCAACATCTGTGGATACAGGACGAGTAACTTTGAGTG 247  
 QY 81 GlyAlaAlaProHisProPheLeuHisArgTyrTyrGlnArgGlnLeuSerSerThrTyr 100  
 DB 248 GGGTGGCCCCCACCCTTCTCTGATCCCTACTACAGAGGAGCTGTCCAGCAATAC 307  
 QY 101 ArgAspLeuArglyGlyValTyrValProTyrThrGlnGlyLyLeuTyrGluGluLeu 120  
 DB 308 CGGAGCTCCGAGAGGGGTGTGTGTGCTTACACCAGGGCAAGTGGAGGGAGCTG 367  
 QY 121 GlyThrAspLeuValSerLeuPheProHisArgTyrProAsnValThrValArgAlaAsnIleAla 140  
 DB 368 GGCACCGACCTGTGTAAAGATCCGCCATGGCCCCCAAGTCACTGTGTGGCAACATGTCT 427  
 QY 141 AlaIleThrGlnSerAspLySerPhePheIleAsnGlySerAsnTyrGluGlyIleGluGly 160  
 DB 428 GCCATACATGATCAAGCAAGTTCTTCAATACAGGCTCCACTCGGAGAGGACCTCTGGGG 487  
 QY 161 LeuAlaTyrAlaGluIleAlaArgProAspAspSerLeuGluProPhePheAspSerLeu 180  
 DB 488 CTGGCTTAATGCTAGATTCGACAGCTGACGACTCCCTGGAGCCTTTCTTGACTCTG 547  
 QY 181 ValIleGlnThrHisValProAsnLeuPheSerLeuGlnLeuCyArglyAlaGlyPhePro 200  
 DB 548 GTAAACAGACCCACGCTTCCCACTCTTCTCCCTGAGCTTGTGTGTGTGCTGCTTCC 607  
 QY 201 LeuAsnGlnSerGluValLeuAlaSerValGlyGlySerMetIleIleGlyGlyIleAsp 220  
 DB 608 CTCACACAGCTGAAGTGGCTGCTGTGCGAGGAGCATGATCATTTGAGAGTATCGAC 667  
 QY 221 HisSerLeuTyrThrGlySerLeuTyrTyrThrProIleArgArgGluTyrTyrGlu 240  
 DB 668 CATTCCTGTACACAGGAGCTCTGTGTATACACCATCCGGCGGAGGTGTATTATGAG 727  
 QY 241 ValIleIleValArgValGluIleAsnGlyGlnAspLeuValMetAspCyArglyGluTyr 260  
 DB 728 GTGATCATTTGTGGGGTGGAGATCAATGACAGAGATCTGAAATATGACATGCAAGAGTAC 787  
 QY 261 AsnTyrAspLySerSerIleValAspSerGlyThrThrAsnLeuArgLeuProLyVal 280  
 DB 788 AACATATGACAGAGCATTTGAGACAGTGGACCAACCACTTGTGTTGCCCAAGAAAGT 847  
 QY 281 PheGluAlaIleValIleSerIleValAlaSerSerThrGluLySerPheProAspGly 300  
 DB 848 TTGGAAGCTCAAGTCAATCATCAAGGAGGCTCTCCACGAGAAATTCCTGATGATG 907  
 QY 301 PheTyrLeuGlyGluGlnLeuValCysTyrGlnAlaGlyThrThrProTyrAsnIlePhe 320

DB 908 TTCTGCTAGAGAGACAGCTGATGTGCGAAGCGGACACACCCTTGGAAATTTTC 967  
 QY 321 ProValIleSerLeuTyrLeuMetGlyGluValThrAsnGlnSerPheArgIleThrIle 340  
 DB 968 CCAGTCACTCACTTACTTAATGGGTAGGTATACCAACAGTCTCTCCGATACCATC 1027  
 QY 341 LeuProGlnGlnTyrLeuArgProValGluAspValAlaThrSerGlnAspAspCysTyr 360  
 DB 1028 CTTCCGACAAATCACTCCGCGCCAGTGAAGATGTGGCCAGCTCCCAAGACAGCTTAT 1087  
 QY 361 LysPheAlaIleSerGlnSerSerThrGlyThrValMetGlyAlaValIleMetGluGly 380  
 DB 1088 AAGTTGCATCTCAAGTATCCACGGGCACTGTATGGAGCTGTATATCATGAGAGGCG 1147  
 QY 381 PheTyrValIleValPheAspArgAlaArgLyValArgIleGlyPheAlaValSerAlaCysHis 400  
 DB 1148 TTCTACGTGTCTTGTGATCGGGCCGAAAGAAATGTGGCTTGTGTGTCATGAGCGCTTGCAT 1207  
 QY 401 ValHisAspGluPheArgThrAlaAlaValGluGlyProPheValThrLeuAspMetGlu 420  
 DB 1208 GTGCAGATGATTCAGAGCGGACGCGGTGAGAGGCCCTTTGTTCATCTTGACATGGAA 1267  
 QY 421 AspCysGlyTyrAsnIleProGlnThrAspGlnSerThr 433  
 DB 1268 GACTGTGGCTACAACTTCCACAGACAGATGATCAACC 1306  
 RESULT 3  
 AY417362 1506 bp DNA linear GSS 17-DEC-2003  
 LOCUS Mus musculus BACE gene, VIRUTAL TRANSCRIPT, partial sequence,  
 DEFINITION genomic survey sequence.  
 ACCESSION AY417362.1 GI:39773322  
 VERSION AY417362.1 GI:39773322  
 KEYWORDS GSS.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE  
 AUTHORS 1 (bases 1 to 1506)  
 Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejaritwal,A.,  
 Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,  
 Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Snihsy,J.J.,  
 Adams,M.D. and Cargill,M.  
 Inferring nonneutral evolution from human-chimp-mouse orthologous  
 gene trios  
 JOURNAL Science 302 (5652), 1960-1963 (2003)  
 PUBMED 14671302  
 REFERENCE  
 AUTHORS 2 (bases 1 to 1506)  
 Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejaritwal,A.,  
 Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,  
 Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Snihsy,J.J.,  
 Adams,M.D. and Cargill,M.  
 Direct Submission  
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
 Rockville, MD 20850, USA  
 COMMENT  
 This sequence was made by sequencing genomic exons and ordering  
 them based on alignment.  
 FEATURES  
 source 1..1506  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 <1..>1506  
 /gene="BACE"  
 /locus\_tag="HCM6198"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 7.3e-240 Length: 1506  
 Score: 2216.00 Matches: 418  
 Percent Similarity: 97.69% Conservative: 5  
 Best Local Similarity: 96.54% Mismatches: 10  
 Query Match: 96.85% Indels: 0

DB: 9 Gaps: 0

US-10-726-967a-84 (1-433) x AY417362 (1-1506)

QY 1 ThrGlnHisGlyIleArgLeuProLeuArgSerGlyLeuGlyAlaProLeuGlyLeu 20

DB 64 ACCCATCTCGGCATCCGGCTGCCCTTCCGACGGCTCGGACGGGCCACCCCTGGGCTG 123

QY 21 GluIleAsnLeuGluThrAspGluProGluGluProGlyArgArgGlySerPheVal 40

DB 124 AGGCTGCCCCGGAGACCGGACGAGGAAATCGAGAGAGCTGGCCGGAGAGCACCTTTGTC 183

QY 41 GluMetValAspAsnLeuArgGlyLysSerGlyGlnGlyTyrTyrValGluMetThrVal 60

DB 184 GAGTGTGTGACAACTGAGGGGAAAGTCCGGCCAGGGCTACTATGTGTGAGATGACCGTA 243

QY 61 GlySerProGlnThrLeuAsnIleLeuValAspThrGlySerSerAsnPheAlaVal 80

DB 244 GCGAGCCCCCAGACGCTCAACATCTGTGTGAGACGGGACGTATGTACTTTGACAGTG 303

QY 81 GlyAlaIleProHisProPheLeuHisArgTyrTyrGlnArgGlnLeuSerSerThrTyr 100

DB 304 GGGGCTGCCCCACACCTTCTTCGTCATCGCTACTACAGAGCAGCTGTCCACACATAT 363

QY 101 ArgAspLeuArgLysGlyValTyrValProTyrThrGlnGlyLysTrpGluGluLeu 120

DB 364 CGAGACCTCCGAAAGGGGTGTATGTCCCTTACACCCAGGGCAAGTGGAGGGGAACTG 423

QY 121 GlyThrAspLeuValSerIleProHisGlyProAsnValThrValArgAlaAsnIleAla 140

DB 424 GGGACGACCTGTGTGACATCCCTCATGGCCCCCAACGTCATGTGGCTGCCAACTTGCT 483

QY 141 AlaIleThrGlySerAspLysPhePheIleAsnGlySerAsnTrpGluGlyIleLeuGly 160

DB 484 GCCATCACTGAAACCGAGCAAGTTCTTCAATCATGATGTTCCAACTGGGAGGCACTCTGA 543

QY 161 LeuAlaTyrAlaGluIleAlaArgProAspAspSerLeuGluProPhePheAspSerLeu 180

DB 544 CTGGCTTATGTCTGAGATTGCCAGGCCGAGCACTCTTTGAGAGCCCTTCTTACCTCCCTG 603

QY 181 ValIleGlnThrHisAlaProAsnLeuPheSerLeuGlnLeuGlyAlaGlyPhePro 200

DB 604 GTGAACCAAGCCCAACATTCCTCAACATCTTCCCTGACCTGTGGCCCTGCTTCCCTCCC 663

QY 201 LeuAsnGlnSerGlyValLeuAlaSerValGlyLysMetIleIleGlyGlyIleAsp 220

DB 664 CTCACACCAAGCAGGCACTGGCTCGTGGAGGAGAGATATCATTTGGTGTATTCAGC 723

QY 221 HisSerLeuTyrThrGlySerLeuTyrTyrThrProIleArgArgGluTyrTyrGlu 240

DB 724 CACTCGCTTATCACCGGCGAGTCTCTGTGTACACACCATCCGGCGGAGTGTATTATGAA 783

QY 241 ValIleIleValArgValGluIleAsnGlyGlnAspLeuLysMetAspCysLysGluTyr 260

DB 784 GTGATCATTTGTACGTGTGAAATTCATATGTCAAAGATCTCAAGATGACTCGCAAGAGATAC 843

QY 261 AsnTyrAspLysSerIleValAspSerGlyThrThrAsnLeuArgLeuProLysVal 280

DB 844 AACTACGACAAAGACATTGTGTGACAGTGGGACCAACCACTTCGTTCCCAAGAAAGTA 903

QY 281 PheGluAlaIleAlaValLysSerIleLysAlaIleAspSerThrGlyLysPheProAspGly 300

DB 904 TTTTGAAGCTGCGGTCAAGTCCATCAAGGACAGCTCTCGACGAGAAATGTCCCGAGTGGC 963

QY 301 PheTrpLeuGlyGluGluLeuValCysTrpGlnAlaGlyTyrThrProTrpAsnIlePhe 320

DB 964 TTTTGGCTAAGGGACACAGCTGTGTGTGGCAAGAGGACAGACCCCTTGGAACATTTTC 1023

QY 321 ProValIleSerLeuTyrLeuMetGlyGluValThrAsnGlnSerPheArgIleThrIle 340

DB 1024 CCAGTCATTTCACCTTACCTCATGGGTGAGATCACCAATCATGCTTCCGATACCACTC 1083

QY 341 LeuProGlnGlnTyrLeuArgProValGluLysValAlaThrSerGlnAspAspCysTyr 360

DB 1084 CTTCCTCAGCAATACCTACGGCCGGTGGAGAGCGGGCCAGCTCCCAAGAGCATGTAC 1143

QY 361 LysPheAlaIleSerGlnSerSerThrGlyThrValMetGlyAlaValIleMetGluGly 380

DB 1144 AGTTGGCTGTCTCAAGTCAATCCAGGGCACTGTATAGGAGCGCTCATCATGAAAGT 1203

QY 381 PheTyrValIlePheAspArgAlaArgLysArgIleGlyPheAlaValSerAlaCysHis 400

DB 1204 TTCTATGTCTGCTTTCATCGACGCGGAAAGCAATTTGGCTTGTCTGTCAAGCGCTTGCCAT 1263

QY 401 ValHisAspGluPheAspThrAlaAlaValGluGlyProPheValThrLeuAspMetGlu 420

DB 1264 GTGACAGATGATTCAGACGCGCGCATGTGAAAGTCCGTTGTTCACGGCAGATCGAA 1323

QY 421 AspCysGlyTyrAsnIleProGlnThrAspGluSerThr 433

DB 1324 GACTGTGCTTCAACATTTCCCAAGACAGATGATCAACA 1362

RESULT 4  
AK041285  
LOCUS  
DEFINITION  
AK041285 3634 bp mRNA linear HTC 03-Apr-2004  
Mus musculus adult male aorta and vein cDNA, RIKEN full-length  
enriched library, clone:A530097B07 product:beta-site App cleaving  
enzyme, full insert sequence.  
ACCESSION  
AK041285  
VERSION  
AK041285.1 GI:26334342  
KEYWORDS  
HTC; CAP trapper.  
SOURCE  
Mus musculus (house mouse)  
ORGANISM  
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE  
AUTHORS  
Carninci, P. and Hayashizaki, Y.  
TITLE  
High-efficiency full-length cDNA cloning  
JOURNALS  
METH. ENZYMOL. 303, 19-44 (1999)  
PUBMED  
99279253  
10349636  
REFERENCE  
AUTHORS  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, K., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.  
TITLE  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes  
JOURNALS  
METH. ENZYMOL. 303, 19-44 (1999)  
PUBMED  
20499374  
11042159  
REFERENCE  
AUTHORS  
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,  
Kono, H., Akiyama, J., Nishi, K., Kikunai, T., Tashiro, H., Itoh, M.,  
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,  
Yamanoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, B., Watanabe, M.,  
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,  
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.  
TITLE  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multiplexed sequencer  
JOURNALS  
METH. ENZYMOL. 303, 19-44 (1999)  
PUBMED  
20530913  
11076861  
REFERENCE  
AUTHORS  
The RIKEN Genome Exploration Research Group Phase II Team and the  
FANTOM Consortium.  
TITLE  
Functional annotation of a full-length mouse cDNA collection  
JOURNALS  
NATURE 409, 685-690 (2001)  
PUBMED  
11076861  
REFERENCE  
AUTHORS  
The PANTOM Consortium and the RIKEN Genome Exploration Research  
Group Phase I & II Team.  
TITLE  
Analysis of the mouse transcriptome based on functional annotation  
of 60,770 full-length cDNAs  
JOURNALS  
NATURE 420, 563-573 (2002)  
PUBMED  
12042159  
11076861  
REFERENCE  
AUTHORS  
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,

Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kankawa, T., Katch, H., Kawai, D., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koye, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Nunazaki, R., Ohno, M., Ohnaka, N., Okazaki, Y., Saito, R., Satoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shideta, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission  
Submitted (16-JUN-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.  
URL: http://fantom.gsc.riken.jp/  
URL: http://genome.gsc.riken.jp/

FEATURES

source

CDS

1..3634  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="PANTOM:DB:A530097B07"  
/db\_xref="taxon:10090"  
/clone="A530097B07"  
/sex="male"  
/tissue\_type="aorta and vein"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="adult"  
450..1955  
/note="unnamed protein product; beta-site APP cleaving enzyme (MGI:MGI:1346542, GB|NM\_011792, evidence: BLASTN, 98%, match=3874)  
putative"  
/protein\_id="BAC30889.1"  
/codon\_start=1  
/db\_xref="GI:26334343"  
/translation="MAPPAHMLWVSGMLPACQTHGRLPLRSGLAPPLGRLP  
RETDESEBGRGSEFVEMVDNIRGSGGTYEMTVGSPPTLNIIVDTGSSNPVAV  
AAPPLHRYORQLSSYRLDKGVYPTQKWEGLSTDLVSI.PHGPNTVPRVNT  
AATTESDKFPINGSNMEGLGLAVAIAPDLSLBPFPGLVYQTHIPNIFSLQCGA  
GFLNQTBALAVSGSMIIGGIDHSLWTTPRREYVYVIVIRVINGQDLM  
DCKEYVYKSDIVSGTNNRLPKVPEAAVKSIIKAASSTKEKDFMGLGOLVCMQAG  
TTPWNI.FPVYSLYLMGEVNTQSPRITILLPQYIRPVEDVATSGDDCYKPAVSSSTGT  
VMGAVIMEGFYVFDRAKRIQFAVSAACHDFRTAAVEGPPVTADMECGNINPOT  
DESTLMTIYVMAICALFMLPLCLMVCCWRCLRHODHDFADDISLKK"

ORIGIN

Alignment Scores:

Score: 2.9e-239 Length: 3634  
Pred. No.: 2216.00 Matches: 418  
Percent Similarity: 97.69% Conservative: 5  
Best Local Similarity: 96.54% Mismatches: 10  
Query Match: 96.85% Indels: 0  
DB: 3 Gaps: 0

US-10-726-967a-84 (1-433) x AK041285 (1-3634)

Qy 1 ThrglnhiaglylleargLeuproLeuargserglyLeuglyglalaProleuglyLeu 20  
Db 513 ACCCATCGGCATCGGCTGCCCTTCGACGGCGCTGGACAGGCCACCCCTGGGCGT 572  
Qy 21 GluileanleugluthrappgluglupProgluglupProglyatargaglysePheval 40  
Db 1653 TTCTATGTCGTCTTCATGAGGCCGGAAGCGAATTGGCTTGGCTGACGCGTCCAT 1712

Db 573 AGGCTGCCCGGAGACCGACGAGATCGAGAGAGCTCGGACCGGAGAGCACTTTGTG 632  
Qy 41 GluuecValaapnaLeuAarglylysergetynginglytyrtyrValaGluuecThrVal 60  
Db 633 GAGATGTGGACCAACTGAGGGGAAGTCGCGCCAGGCTCTATGTGAATGACCGTA 692  
Qy 61 GlyserProProglInthrLeuamilleuValaApthrGlyserSerAsnPhelaVal 80  
Db 693 GGCACCCCCCAGACACCTCAACATCTGGTGGACAGCGGCAAGTAACTTTCGACATG 752  
Qy 81 GlyalaAaprohisAproPheleuuhisaargtyrtyrGlnaargInLeuSerSerThyr 100  
Db 753 GGGGCTGCCACACCTTTCTGTCATCTACAGAGGACGCTGCCAGACATAT 812  
Qy 101 ArgapLeuAaglyglyValaTyraProtyrThringInglylvetrGluIngluLeu 120  
Db 813 CGAGACCTCCGAGAGGGGTGTATGTGCTTACACCCAGGGGCAAGTGAAGGGGAACG 872  
Qy 121 GlyThrAapLeuValSerileProhisAprohisAprohisAprohisAprohisA 140  
Db 873 GGCACCGACCTGGTGAACATCTCTGATGAGCCCAACGTCATCTGCTGCCAATTCGT 932  
Qy 141 AlailethrGluSerAaplyAepheleleanglyserAanttrpGluGlyleugly 160  
Db 933 GCCATCATGATCGACAGCAAGTTTCATCATGTTCCAACTGGAGGGCATCTAGGG 992  
Qy 161 LeuAaTyraAagluilleaAargProAapSerleuGluProPheAapSerleu 180  
Db 993 CTGGCTATGCTGAGATGCGAGCGCCGACGACTTTGGAGCCCTTCTTGACCTCCG 1052  
Qy 181 VallyseGlnThhisValProamLeupheserleuGlnLeuCyglValaGlyPhePro 200  
Db 1053 GGAAGACGACCCAGCAATCCCAACATCTTTCTCGACGCTGTGGCCCTGCTTCCC 1112  
Qy 201 LeuAangInserGluValaLeuAaSerValaGlylyserMelleileglylyleap 220  
Db 1113 CTCACGACGACCGAGGACCTGGCTCGGAGGAGGAGATGATGATGATGATGATGAC 1172  
Qy 221 HisserleuTyrThglyserleuTrpTyrrProilleAargAargluTrpTyrrGlu 240  
Db 1173 CACTGCTATACACGGCGAGCTCTCTGATGACACACCATCCGGGAGAGGTATATGAA 1232  
Qy 241 ValilleleValaArgValaGluilleaanglyGlnaapLeuValMetAapCylyseGluTy 260  
Db 1233 GGATCATATGACGATGAGAAATCAATGCTCAAGATCTCAATGATGATGATGATGATG 1292  
Qy 261 AantTyraAaplySerileValaAapSerGlyThrThrasnLeuAargLeuProlylyVal 280  
Db 1293 AACTACGACAAAGACATTTGTGACAGTGGGACCAACCACTTCGCTGCCAAGAAAGTA 1352  
Qy 281 PheGluAaAaValaIysSerilelysaAaAaSerSerThGluAlyPheProAapGly 300  
Db 1353 TTTGAAGCTGCGCTCAAGTCAATCAAGGCAAGCTCTCTCAAGAGAAAGTCCCGATGCG 1412  
Qy 301 PheTrpLeuGlyGluGlnLeuValCystrpGlnaAaglyThrThrProTrpAsnIlePhe 320  
Db 1413 TTTTGGCTAGGGAGAGAGCTGATGTGTGTGGCAAGGAGGACCAACCTTGGAAATTTTC 1472  
Qy 321 ProvalileSerleuTyreleuMetGlyGluValaThrasnGlnSerPheAargIleThrIle 340  
Db 1473 CCAGTCATTTCACTTCACTCAATGGTGAAGTCAACCAATCAAGTCTTCCGATCACATTC 1532  
Qy 341 LeuProGlnGlnTyreleuAargProValaGluAapValaAaThrSerGlnaAapCystr 360  
Db 1533 CTTCCTCAACAATCACTTACGCGCGGTGGAGAGCTGGCCACGCTCCCAACAGCATGTTTC 1592  
Qy 361 LysPheAaAaIleSerGlnSerSerThrArgIleValaMetGlyValaValIleMetGly 380  
Db 1593 AAGTTGGCTGTCAAGTATCAACGAGGACATGTTATGAGACCTCATCATGAGAGGT 1652  
Qy 381 PheTyrValaValaPheAapAargAaarglyAargIleGlyPheAaValaSerAaCyhis 400  
Db 1653 TTCTATGTCGTCTTCATGAGGCCGGAAGCGAATTGGCTTGGCTGACGCGTCCAT 1712

QY 401 ValHSAAGGluPheGlyThrAlaAlaValGluGlyProPheValThrLeuAspMetGlu 420  
|||||  
Db 1713 GTGACGAGTGGATTGAGGACGGCGAGTGGAGGTCCGTTTGTACGGGACATGGAA 1772  
QY 421 AAGCGAGGlyTyrAsnIleProGlnThrAspGluSerThr 433  
|||||  
Db 1773 GAGTGTGGCTACAAACATTCCCGACAGATGATGACAA 1811  
RESULT 5  
AK014464  
LOCUS 3859 bp mRNA linear HTC 03-Apr-2004  
DEFINITION Mus musculus 16 days embryo head cDNA, RIKEN full-length enriched  
library, clone:4122401C04 product:beta-site APP cleaving enzyme,  
full insert sequence.  
ACCESSION AK014464  
VERSION AK014464.1 GI:12852334  
KEYWORDS HTC; CAP trapper.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Mammalia; Euteleostomi; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Euteleostomi; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE  
AUTHORS Carninci, P. and Hayashizaki, Y.  
TITLE High-efficiency full-length cDNA cloning  
JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
MEDLINE 99279253  
PUBMED 10349636  
REFERENCE  
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes  
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)  
MEDLINE 20499374  
PUBMED 11042159  
REFERENCE  
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P.,  
Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,  
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,  
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,  
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunura, S., Kawai, J.,  
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
TITLE RIKEN integrated sequence analysis (RISA) system-384-format  
sequencing pipeline with 384 multicapillary sequencer  
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)  
MEDLINE 20530913  
PUBMED 11076861  
REFERENCE  
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the  
FANTOM Consortium.  
TITLE Functional annotation of a full-length mouse cDNA collection  
JOURNAL Nature 409, 685-690 (2001)  
REFERENCE  
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research  
Group Phase I & II Team.  
TITLE Analysis of the mouse transcriptome based on functional annotation  
of 60,770 full-length cDNAs  
JOURNAL Nature 420, 563-573 (2002)  
REFERENCE  
AUTHORS Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H.,  
Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,  
Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,  
Hirooka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,  
Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Konda, M.,  
Koyama, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K.,  
Nomura, K., Numasaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,  
Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,  
Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y.,  
Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T.,  
Tejima, Y., Toyama, T., Yamamura, T., Yasunishi, A., Yoshida, K.,  
Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

TITLE  
JOURNAL  
COMMENT  
Direct Submission  
Submitted (10-JUN-2000) Yoshihide Hayashizaki. The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Science Center (GSC),  
RIKEN Yokohama Institute; 1-7-22 Sueniwa-cho, Tsunumi-ku, Yokohama,  
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp,  
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,  
Fax: 81-45-503-9216)  
Please visit our web site (http://genome.gsc.riken.jp/) for further  
details.  
cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Science Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues. First strand cDNA was primed with a primer  
[5' GAGAGAGAGAGATCCAGAGATCCGCTTTTCTTTTCTTTT 3'], cDNA was  
prepared by using trehalose thermo-activated reverse transcriptase  
and subsequently enriched for full-length by cap-trapper. Second  
strand cDNA was prepared with the primer adapter of sequence [5'  
GAGAGAGAGATCCGAGATTAAATTAATTAATCCCGCCCGCC 3']. cDNA was cleaved  
with BamHI and XhoI. cDNA of size comprised longer than 7 kb was  
selected before cloning. Vector: a modified pBluescript KS(+) after  
bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI, 3'  
end: BamHI. Host: DH10B.  
FEATURES  
source  
Location/Qualifiers  
1..3859  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM DB:4122401C04"  
/db\_xref="taxon:10090"  
/clone="4122401C04"  
/clone="4122401C04"  
/issue\_type="head"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="16 days embryo"  
430..1935  
/note="unnamed protein product; beta-site APP cleaving  
enzyme (MGI:1346542, GB|NM\_011792, evidence: BLASTN,  
98%, match=3874)  
putative"  
/protein\_id="BAB29370.1"  
/db\_xref="GI:12852335"  
/translation="MAAPALHWLWVSGMLPAQGNLGRPLPLSGLAGPPLGRPL  
RPPDSSEPRGRGSGVWVDNRKSGSGGTIVEMTGSPPQTLLIVDGSNRAVG  
AAHPFLRYQRQUSSTYRDLKGVLYVYTGKWEGLGDLVSIPIHPNPTVANI  
AALTDESDFPINSNMEGILGLAYELIAPDDSLFEPDLVKQTHIPVIFSLQCSGA  
GFLNQTEALASVSGSMIIGIDHSLYTGSLWYTPDIRREYEVILIVRGEIOLDKM  
DKEKNYDKSIYDSTGNLRLLPKVPEAAVKSIKAASSTREKPFMFGIOLWQAG  
TTPMNIFFPVISLYIMGEVTNOSFRITLIPQOVLREVDVATSDODCKYPAVQSSTGT  
VMGAVIMGEFVVPEAPRRIRIGPAVSACVHDEPRTAAVEGFPYADMDGCSNIPQT  
DESLMTIAYVMAALCALFMDPLCLMVCQMRCLRHQHDHFDADISLKL"  
ORIGIN  
Alignment Scores:  
Pred. No.: 3,19e-239 Length: 3859  
Score: 2216.00 Matches: 418  
Percent Similarity: 97.69% Conservative: 5  
Best Local Similarity: 96.54% Mismatch: 10  
Query Match: 96.85% Indels: 0  
Gaps: 0  
US-10-726-967A-84 (1-433) x AK014464 (1-3859)  
QY 1 ThrGlnHicGlyTlleArgLeuProLeuArgSerGlyLeuGlyAlaProLeuGlyLeu 20  
|||||  
Db 493 ACCCATCTGGGATCCGCTGCCCTTCGCACAGCGCTGGCGAGGCCACCCCTGGGCGCTG 552  
QY 21 GluileAsnLeuGlyThrAspGluPheGlyProGluGluGlyPheArgGlySerPheVal 40  
|||||  
Db 553 AAGCTGCCCCGGGAGACCGACAGAAATCGAGAGAGCCCTGGCGGAGGCGAGCTTTGTG 612  
QY 41 GluMetValAspAsnLeuArgGlyLysSerGlyGlnGlyTyrTyrValGluMetThrVal 60  
|||||

Db	613	GAGAGTGGTGAACCACTGAGGGGAAAGTCGGCCAGGGCTACTATGTGAGATGACCCTA	672
Oy	61	GIYSEPRPFGINTThrLeuasnilleuValAspThrGIYSESErAsnPhelAval	80
Db	673	GGGAGCCCCCAGACGCTCAACATCTGGTGGACAGGGGAGTAACTTGGCACTG	732
Oy	81	GIYAlaAlaProHiasProPheleuHiasrGIYrYrGlnArgGlnLeuSErThrYr	100
Db	723	GGGGGTGGCCACACCTTTCCTGCATGCTACTACAGAGGACAGCTGTCCAGCACATAT	792
Oy	101	ArgAspLeuAArglyValTyrtValProTYrThrngInlytRTrgInlytRgInleu	120
Db	793	CGAGACCTCCGAAAGGTGTATGTGCCCTACACCCAGGGGCAATGGGAGGGGAACTG	852
Oy	121	GIYThrAspLeuValSerilleProHiasGlyProAsnValThrValArgAlaasnillea	140
Db	853	GGCACCGACCTGGTGAAGCATCCCTCATGGCCCCAACGTCACTGTGGCGCAACTTCT	912
Oy	141	AlaIleThrGlySErAspLyAspPheHieleaenglySErAsnTrpGlnlytleuGly	160
Db	913	GGCATCACTGAATCGGACAAAGTTCTTCAATAGTGTCCCACTGGAGGGGCATCTTAAGG	972
Oy	161	LeuAlaTYrAlaGlnIleAlaArgProAspAspSerleuGlnProPhaPheAspSerleu	180
Db	973	CTGGCGCTATCGTAATTCGACAAATTCGAGGCCACGACTTTTGGACCCCTTCTTGAATCCCTG	1033
Oy	181	VallyrSGINTThHieValProAsnIlePheSErLeuGlnLeuCYGlyValaglyPhePro	200
Db	1033	GTGAAGACGACCCACCACTTCCCAACATCTTTTCCCTGACCTGTGGCGCTGGCTTCCC	1092
Oy	201	LeuAsnGlnSErGlyValleuAlaSErValGlylySErMetIleIleGlylylleaP	220
Db	1093	CTCAACCAAGACCGAGCACTGGCTCGGTGGGAGGAGCATGATGATGGTGTGATCGAC	1152
Oy	221	HISerLeuTYrThGlySErleuTrpTYrThrProIleArgArgGlnTrpTYrTYrGln	240
Db	1153	CACCTGCTATACCGGGCAGTCTTCGTACACACCATCCGGGGAGGTGATTAATGA	1212
Oy	241	ValIleIleIleValArgValGlnIleAsnGlyIleAspLeuLYMetAspCYblyrSGlnTYr	260
Db	1213	GTGATCATTTGTACGTGTGGAAATTCANGTGTCAAGATCTCAAGATGGATCTGCAAGAGATNC	1272
Oy	261	AsnTYrAspLySErIleValAspSErGIYThrThrAsnLeuArgLeuProIlyAlaVal	280
Db	1273	AACThGACAAAGAGCATTTGTGACAGTGGGACCAACCAACTTCGCTTGGCCAAAGAAATGA	1332
Oy	281	PheGlnAlaAlaValIlySErIlelyValAlaAsEserThrGlnLYPheProAspGly	300
Db	1333	TTTGAAGCTGCCGTCAAGTCATCAAGGACGCTCTCCACGAGAAAGTTCCCGAATGGC	1392
Oy	301	PheTrpLeuGIYGlnGlnleuValCYeTrpGlnAlaGlyThrThrProTrpAsnIlePhe	320
Db	1393	TTTTGGCTAAGGAGAGCTGTGTGTGTGGCAAGGACGACCAACCCCTTGGAACTTTTC	1452
Oy	321	ProValIleSErleuTYrleuMetGIYGlyValThrAsnGlnSErPheArgIleThrIle	340
Db	1453	CCAGTCATTTCACTTACTCTCATGGGTGAAGTCAACAACTAGTCCCTTGGCATCACATC	1512
Oy	341	LeuPTrGlnGlnTYrleuArgProValGluAspValAlaThrSErGlnAspAspCYrTYr	360
Db	1513	CTTCCCTCAACAATTAACCTAGCGCGGTGGAGGACGTGGCCACGTCCCAACACACACTGTAC	1572
Oy	361	LySPheAlaIleSErGlnSErThrnglyThrValMetGIYAlaValIleMetGlnGly	380
Db	1573	AAGTTGCGTGTCCACAGTCATCCACGGGACCTGTTATGGAGACCGTCATCTGAAGGT	1632
Oy	381	PheTYrValValPheAspArgAlaArglybArgIleGlyPheAlaValSerAlaCYbHias	400
Db	1633	TTCTATGTGTCTTTCATTCGAGCCGAAAGCAATTGGCTTTGTGTGTGACGCGTTGCCAT	1692
Oy	401	ValHiaspGlyPheArgThrAlaAlaValGlyGlyProPheValThrleuAspMetGln	420

DB	1693	GTGACGATGATGTTACAGGACGCGGAGTGAGMAGTCCGTTTGTTACGCGACAGATGAGA	1752
QY	421	AspCysgltYTYraanlleProGlnThAspGlnSerThr	433
Db	1753	GACTGTGGCTACCAACATTCCCGACAGATGATGACAA	1791
RESULT 6			
AK033112			
LOCUS		3877 bp	mRNA
DEFINITION		Mus musculus 15 days embryo male testis cDNA, RIKEN full-length enriched library, clone:8030431G04 product:beta-site APP cleaving enzyme, full insert sequence.	HTC 03-APR-2004
AK033112			
ACCESSION		AK033112.1	GI:26328834
VERSION		HTC; CAP trapper.	
KEYWORDS		Mus musculus (house mouse)	
ORGANISM		Mus musculus	
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			



Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
Kanagawa 230-0045, Japan (E-mail: genome-res@gs.c.riken.jp,  
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,  
Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in Riken.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.  
Please visit our web site for further details.  
URL: http://genome.gsc.riken.jp/  
URL: http://fantom.gsc.riken.jp/.

## FEATURES

## Source

1. 3877

Location/Qualifiers

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C57BL/6J"

/db\_xref="FANTOM, DB:8030431G04"

/db\_xref="taxon:10090"

/clone="8030431G04"

/sex="male"

/tissue\_type="testis"

/clone\_lib="RIKEN full-length enriched mouse cDNA library"

/dev\_stage="15 days embryo"

/note="unnamed protein product; beta-site APP cleaving  
enzyme (MGI:MGI:1346542, GB|NM\_011792, evidence: BLASTN,  
98%, match=3874)

putative"

/codon\_start=1

/protein\_id="BAC28156.1"

/db\_xref="GI:26328835"

/translation="MAPALHLLWVSGMLPAQGNHILGIRLIRSLGAGPRLGRP  
REDDSEBGRGRGFVEMVDNLRGSGGYVEMWVGSPPPTLNIIVDTGSSNRAVG  
AAPRLHRYRQRLSTYRDLARKYVYTYGKEGLDVLVSIPIHGNVYVANI  
AATTESDKRFINSNWEGLGLAVAIAPDLSLEFDSLVKQTHIPNIFSLQCGA  
GPELQNTALASVSGSMIGIDHSLYTSLSMTPIRRREYEVIIIVEIINGQDLK  
DCEVNYDSIVDSGTNNLRLPKVFEAAVKSIIKAASTKEKPDGFMLEQULVCWQAG  
TTPMNIPIVLSLMEGTNOSFRITILPQOYLRPVEDVATSDQDCYKPAVSOSTGT  
VNGAVIMEGFYVFPDARKRIGFVASACHVDEPRFAVAGSPFVADMDKCYRNIPOT  
DESTLMTIYVMAAICALPMLPLCLMVCWRCRLRHRQHDHDFADISLKL"

## ORIGIN

## Alignment Scores:

Pred. No.: 3, 21e-239 Length: 3877  
Score: 2216.00 Matches: 418  
Percent Similarity: 97.69% Conservative: 5  
Best Local Similarity: 96.54% Mismatches: 10  
Query Match: 96.85% Indels: 0  
DB: 3 Gaps: 0

US-10-726-967a-84 (1-433) x AK031112 (1-3877)

QY 1 ThrGlnHisGlyIleArgLeuProLeuArgSerGlyLeuGlyGlyAlaProLeuGlyLeu 20  
DB 513 ACCCATCTCGGACATCGGCTGCCCTTCGCACAGCGGCTGGCAGGGCCACCCCTGGGCTG 572  
QY 21 GlnIleAsnLeuGluThrAspGluGluProGluGluProGluYArgArgGlySerPheVal 40  
DB 573 AGGCTGCGCCCGGAGAGCGACGAGAAATCGAGAGAGCTGCGGAGAGGACGCTTTGTG 632  
QY 41 GlnMetValAspAsnLeuArgGlySerGlyGlnGlyTyrTyrValGlnMetThrVal 60  
DB 633 GAGATGGTGGACAACTGAGGGGAAAGTCCGGCCAGGGCTACTATGTGAGATGACCGCTA 692  
QY 61 GlySerProGluThrIleuAsnIleuValAspThrGlySerSerAspAsnAlaVal 80  
DB 693 GGCACGCCCCCAGACGCTCAACATCTGTGTGACAGGGGAGAGTAACTTTGACGTG 752  
QY 81 GlyAlaAlaProHisProPheLeuHisArgTyrTyrGlnArgGlnLeuSerSerThrTyr 100  
DB 753 GGGGCTGCCCAACACCTTCTCTGATCGTACTACGAGAGGAGCGAGTGTCCAGCAACATAT 812

QY 101 ArgAspLeuArgGlyValTyrValProTyrThrGlnGlyLeuTyrGluGlyLeu 120  
DB 813 CCAAGACCTCCGAAAGAGGTGTGATGTGCTTCAACCCAGGCGCACTGGAGGGGAACTG 872  
QY 121 GlyThrAspLeuValSerIleProHisGlyProAsnValThrValArgAlaAsnIleAla 140  
DB 873 GGCACCGACCTGGTAGATCCCTCATGAGCCCAACGTCATGGCGTGGCCAACTTGCT 932  
QY 141 AlaIleThrGluSerAspGlyPhePheIleAsnGlySerAsnTyrGluIleLeuGly 160  
DB 933 GGCATCATGAAATCGACAAATGTTCTTCAATCAATGTTCAATGAGGGATCTTGA 992  
QY 161 LeuAlaTyrAlaGluIleAlaArgProAspAspSerLeuGluProPheAspSerLeu 180  
DB 993 CTGGCTTATGCTGAGATTTGCAGAGCCGACGACACTCTTGGAGCCCTTCTTGAATCCCTG 1052  
QY 181 ValIysGlnThrHisValProAsnLeuPheSerLeuGlnLeuGlyAlaGlyPhePro 200  
DB 1053 GTGAAGCAGACCCACATTCACCAACATCTTTCCTGACGCTGTGGCGCTGGCTCC 1112  
QY 201 LeuAsnGlnSerGluValIleuAlaSerValGlyGlySerMetIleIleGlyGlyIleAsp 220  
DB 1113 CTCAACGACGACGAGGACCTGCTCGGTGGAGGAGCATGATCATTTGGTATTCAC 1172  
QY 221 HisSerLeuTyrThrGlySerLeuTyrTyrThrProIleArgArgGluTyrTyrGlu 240  
DB 1173 CACTGCTTATACAGGGGAGCTCTCTGTACACACCCATCCGGCGGAGTGTATATGAA 1232  
QY 241 ValIleIleValArgValGluIleAsnGlyValAspLeuMetAspCysLeuGlyTyr 260  
DB 1233 GTGATCATTTGACTGTGGAATATCATGTCAGATTTCAAGATGACGTGCAAGAGTAC 1292  
QY 261 AsnTyrAspIysSerIleValAspSerGlyTyrThrAsnLeuArgLeuProIleVal 280  
DB 1293 AACTACACAAAGACATTTGAGACAGTGGAGACCAACCTTGCCTTCCCAAGAAATG 1352  
QY 281 PheGluAlaIleValIysSerIleValAlaIleSerSerThrCysIlePheProAspGly 300  
DB 1353 TTTAACTGTGCTCAAGTCAATCAACAGGCGCTCTCGACGGAAGTCCCGGATGGC 1412  
QY 301 PheTyrLeuGlyGluGlnLeuValCysTyrGlnAlaGlyTyrThrProTyrAsnIlePhe 320  
DB 1413 TTTTGGCTAAGGGAGGAGCGTGTGTGCTGGCAAGGACGACCCCTTGGAAATTTTC 1472  
QY 321 ProValIleSerLeuTyrIleuMetGlyGluValThrAsnGlnSerPheArgIleThrIle 340  
DB 1473 CCAATCAATTCATTAACCTCATGGGTGAAGTCAACCAATCAGTCTCCGCATCACATC 1532  
QY 341 LeuProGlnGlnTyrIleuArgProValGluAspValAlaThrSerGlnAspAspCysTyr 360  
DB 1533 CTTCCTCAGCAATCACTTACGCGCGGTGAGAGAGAGTGGCCAGCTCCCAACACATGTTTC 1592  
QY 361 LysPheAlaIleSerGlnSerSerThrGlyTyrValMetGlyAlaValIleMetGluGly 380  
DB 1593 AAGTTCGCTGTCAAGTATCCACAGGCGACGTTATGGAGCCGTATCATGGAGGT 1652  
QY 381 PheTyrValValPheAspArgAlaArgIleValGlyPheAlaValSerAlaCysHis 400  
DB 1653 TTTTATGTGTCTTTCATGTGAGGCGGAGCCGAAAGCAATTTGGCTTGTGACGCCCTTGCAAT 1712  
QY 401 ValHisAspGluPheArgThrAlaAlaValGluGlyProPheValThrLeuAspMetGlu 420  
DB 1713 GTGACGATGATGTTCAAGAGCGGAGAGTGAAGGTCCGTTTGTACCGGACATGGA 1772  
QY 421 AspCysGlyTyrAsnIleProGlnThrAspGluSerThr 433  
DB 1773 GACTGTGCTTACAACTTCCCAAGAGATGATCAACA 1811

RESULT 7  
AK049626 4046 bp mRNA linear HFC 03-APR-2004  
LOCUS AK049626  
DEFINITION Mus musculus 12 days embryo spinal cord cDNA, RIKEN full-length



enriched library, clone: C530008K17 product: beta-site APP cleaving enzyme, full insert sequence.

ACCESSION  
AK049626  
VERSION  
AK049626.1 GI:26340361  
KEYWORDS  
HTC; CAP trapper.  
SOURCE  
Mus musculus (house mouse)  
ORGANISM  
Mus musculus

Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
AUTHORS  
1 Carninci, P. and Hayashizaki, Y.  
TITLE  
High-efficiency full-length cDNA cloning  
JOURNAL  
Meth. Enzymol. 303, 19-44 (1999)  
MEDLINE  
99279253  
PUBMED  
10349636

REFERENCE  
AUTHORS  
2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
TITLE  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
JOURNAL  
Genome Res. 10 (10), 1617-1630 (2000)  
MEDLINE  
20499374  
PUBMED  
11042159

REFERENCE  
AUTHORS  
3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kitsuami, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, H., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujikake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matsubiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
TITLE  
RIKEN integrated sequence analysis (RISA) system -384-Format  
JOURNAL  
Genome Res. 10 (11), 1757-1771 (2000)  
MEDLINE  
20530913  
PUBMED  
11076861

REFERENCE  
AUTHORS  
4 The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium.  
TITLE  
Functional annotation of a full-length mouse cDNA collection  
JOURNAL  
Nature 409, 685-690 (2001)  
MEDLINE  
11076861

REFERENCE  
AUTHORS  
5 The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
TITLE  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
JOURNAL  
Nature 420, 563-573 (2002)  
MEDLINE  
11076861

REFERENCE  
AUTHORS  
6 (bases 1 to 4046)  
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Konda, M., Koyama, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.  
TITLE  
Direct SubMISSION  
JOURNAL  
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT  
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

#### FEATURES

source  
1. 4046  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="PANTOM\_DB:C530008K17"  
/db\_xref="taxon:10090"  
/clone="C530008K17"  
/tissue\_type="spinal cord"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="12 days embryo"  
433..1938  
/note="unnamed protein product, beta-site APP cleaving enzyme (MDJ|MG1:1346542, GI|NM\_011792, evidence: BLASTN, 98%, match=3874)  
putative"  
/codon\_start=1  
/protein\_id="AAC33844.1"  
/db\_xref="GI:26340362"  
/translation="MAPALHWLLWVSGMLPAQCTHILGRLPLRSLGAPPLGLRLP  
RETDESEPERGRGSFVEMVDNLKSGSGGYVEMTTGSPQTLNIVDTGSSNFAVG  
AAPDFLHRYVOROLSTYRDLRKGVYVYQKWEGLGTLVSIIPGPNVYVANI  
AATRESDEPFNGSWEGILGLAVAIARPDDSLPEPDSLKOTHIINIFSLQCGA  
GPELQUTALASVGSMTIGTIDSLTYSGLSYMTPIREYEVYIVVEINQDLKM  
DCKEYVYKSLVDSTNLRLPKVFEVAVSIKASSTKPPDFMGLVQVQAG  
TTPNIVFPIVLVLMGEVNTQSFRTTILPQYLRVEAVANSODCYFAVSQSTGT  
VMGAVIMGEFVYVPEPRARKRIGFAVASCHVDFEFAVEGPFVADMEDCYNIPOT  
DSTLMTIAYVMAIICALFMLPLCMVCOMRCRLCRHQHDFADDISLTK"

#### CDS

ORIGIN  
Alignment Scores:  
Pred. No.: 4,466-239 Length: 4046  
Score: 2215.00 Matches: 417  
Percent Similarity: 97.69% Conservative: 6  
Best Local Similarity: 96.30% Mismatches: 10  
Query Match: 96.81% Indels: 0  
DB: 3 Gaps: 0  
US-10-726-967a-84 (1-433) x AK049626 (1-4046)

QY 1 TTTGTHHISGLYITLARGLEUProleuArgSerGlyLeuGlyValAProleuGlyLeu 20  
DB 496 ACCCATCTCGGATCCGCTGCTCCCTTCGACAGGCGCTGGCAGGCGCACCCCTGGGCTTG 555  
QY 21 GTTLeuLeuGlyLThAspGluGluProGluGluProGluYArgArgGlySerPheVal 40  
DB 556 AGCTGCCCCCGGAGACCGACGAGATCGAGAGGCTTGGCCGAGAGGCGAGCTTTGG 615  
QY 41 GtmetValAspAsnLeuArgGlyLysSerGlyGlyGlyTyrTyrValGtmetThrVal 60  
DB 616 GAGATGGTGAACAACCTGAGGGGAAAGTCGGCCAGGCGCTACTATGTGAGATGACATTA 675  
QY 61 GlySerProProGlnThrLeuAsnLleLeuValAspThrGlySerSerAsnPheLysVal 80  
DB 676 GCGAGCGCCCCCAGACGCTCAACATCTCGTGGACACGGGACGAGTAGTAACCTTTGCAAGT 735  
QY 81 GtValAlaProHisProPheLeuHisArgTyrTyrGlnArgGlnLeuSerSerThrTyr 100  
DB 736 GGGGCTGCCCCACACCCCTTCTCGATCGCTACTACAGAGGCAAGCTGCCAGCACATAT 795  
QY 101 ATGAspLeuArgLysGlyValTyrValProTyrThrGlyGlyLysTyrGluGlyLeu 120  
DB 796 CGAGACCTCGAAGAGGTGTGATGTGCTTACACCCAGGCGCAAGTGGAGGGGGAACCTG 855  
QY 121 GtYThAspLeuValSerLleProHisGlyProAsnValThValArgAlaAsnLleAla 140  
DB 856 GGCACGACCTGTGTGCAATCCCTCAATGCCCCCAACGTCATCTGTGTGTCACATGCT 915  
QY 141 AlaLleThGluSerAspLysPhePheLleAsnGlySerAsnTyrGluGlyLleuGly 160

Db	916	GCACATCACTGAATCGGACAGATTCTTCATCAATGTTTCCAACTGGGAGGCATCTCTAGG	975
Qy	161	LeuAlaLeuYrYrIaGluIleAlaAaGrProAaPaSerLeuGluProPheAaPaSerLeu	180
Db	976	CTGGCCATATCTAGATTCAGAGCCGACGACATCTTTTGAGACCCCTTCTTGACCTCCGTG	10359
Qy	181	ValIleGlnThrHisValProAaLeuPheSerLeuGlnLeuCyGluValaGlyPhePro	200
Db	1036	GTGAGAGACAGCCACATTCOCACATCTTTCCCTCGACCTCTGTGGCGCTTGGCTTCCCC	10995
Qy	201	LeuAaGlnSerGluValLeuAlaSerValGlyGlySerMetIleIleGlyGlyIleAaP	220
Db	1096	CTCAACCCAGCCAGGACACTGGCTCGTGGGAGGAGCATGATCATTTGATGATCAAC	1155
Qy	221	HisSerLeuYrThrGlySerLeuYrPyrTrpThrProIleArgArgGluTrpTyrTrpGlu	240
Db	1156	CATCTCCGCTATACCGGACAGCTCTCTGTATACACCCATCTGGCGGAGATGTATTATGAA	1215
Qy	241	ValIleIleValArgValGluIleAaGlnAaPheLeuLysMetAaPheCyLeuGlyIleTyr	260
Db	1216	GTGATCATTTGATCGTGTGAAATTCATATGCTCAAGATCTCAAGATGCACTGCAAGAGATAC	1275
Qy	261	AaYrYrAaPheLysSerIleValAaPaSerGlyYrThrAaLeuAaGluLeuProGlyVala	280
Db	1276	AATCTACGACAAAGAGCATTTGTGACAGTGGACACACCACTTCGCTTCCCAAGAAAGTA	1335
Qy	281	PheGluAlaAlaValIleSerIleLeuYsAlaAaSerSerThrGlnGluLysPheProAaPgly	300
Db	1336	TTTGAAAGCTGCCGTCAAGTCCATCAAGCAGCCTCTCTGACGAGAAAGTTCCCGATGCG	1395
Qy	301	PheTrpPheGlyGluGlnLeuValCysTrpGlnAlaGlyThrProTrpAaIlePhe	320
Db	1396	TTTTGGCTAAGGGAGACAGCTGTGTGTCTGGCAAGCAGCAGACCCCTTGGAAATTTTC	1455
Qy	321	ProValIleSerLeuYrLeuMetGlyGluValThrAaGlnSerPheArgIleIleThrIle	340
Db	1456	CCAGTCATTTCACTTACTCTCAATGGGTAAATGACCAATCAAGTCTTCCGATCACCATC	1515
Qy	341	LeuProGlnGlnTyrIleuArgProValGluAaPheValaIleThrSerGlnAaPaPheCysTyr	360
Db	1516	CTTCCCTCAGCATATACCTACGGCGGGTGGAGACGAGGCACTGCCAAGACGACTGTAC	1575
Qy	361	LysPheAlaIleSerGlnSerSerThrGlyYrValMetGlyValaValIleMetGluGly	380
Db	1576	AAGTTGCGTGTCTCAAGCTATCCACGGGCACTGTATTGGAGCCGTCATCATTAAGAAAGT	1635
Qy	381	PheYrValValPheAaPaArgAlaArgIysArgIleGlyPheAlaValSerAlaCysHis	400
Db	1636	TTCTATGTCGCTTTCGATCGAGCCGAAAGGAAATTTGGCTTGTGTCAAGCTTTCACAT	1695
Qy	401	ValHisAaPglyPheArgThrAlaAlaValGluGlyProPheValThrLeuAaPaMetGlu	420
Db	1696	GTGCACGATGAGTTTCAGAGACGGCGGAGGTGGAAAGTCCGTTTGTATCGGCAGACATGA	1755
Qy	421	AaPheGlyIleYrAaIlePheProGlnThrAaPglySerThr	433
Db	1756	GACTGTGCTCAACATTTCCCAACAGATGAGTCACAA	1794
RESULT 8			
AK046175			
LOCUS	4101 bp	mRNA	linear
DEFINITION	Mus musculus adult male corpora quadrigemina cDNA, RIKEN		HTC 03-APR-2004
VERSION	full-length enriched library, clone: B203646M13		product: beta-site
KEYWORDS	App cleaving enzyme, full insert sequence.		
ACCESSION	AK046175.1	GI:26337868	
VERSION	AK046175.1	GI:26337868	
KEYWORDS	HTC, CAP trapper.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
	Carninci, P. and Hayashizaki, Y.		

TITLE	High-efficiency full-length cDNA cloning
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)
MEDLINE	99279253
PUBMED	10349636
REFERENCE	
AUTHORS	2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE	20499374
PUBMED	11042159
REFERENCE	
AUTHORS	3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Komno, H., Akiyama, J., Nishi, K., Kitenaka, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kaishiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Iwawa, M., Ohara, E., Matsubiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipipillary sequencer
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE	20530913
PUBMED	11076861
REFERENCE	
AUTHORS	4
TITLE	The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium.
JOURNAL	Functional annotation of a full-length mouse cDNA collection
MEDLINE	Nature 409, 685-690 (2001)
PUBMED	5
REFERENCE	
AUTHORS	The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)
MEDLINE	6 (bases 1 to 410)
PUBMED	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Horii, F., Imocanti, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kori, H., Kawai, J., Kojima, Y., Kondo, S., Komno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
REFERENCE	
AUTHORS	Direct Submission
TITLE	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
JOURNAL	
MEDLINE	
PUBMED	
REFERENCE	
AUTHORS	
TITLE	cDNA library was prepared and sequenced in Mouse Genome Encyclopaedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
JOURNAL	Please visit our web site for further details.
MEDLINE	URL: http://genome.gsc.riken.jp/
PUBMED	URL: http://Fantom.gsc.riken.jp/
REFERENCE	Location/Qualifiers
AUTHORS	1. 4101
TITLE	/organism="Mus musculus"
JOURNAL	/mol_type="mRNA"
MEDLINE	/strain="C57BL/6J"
PUBMED	/db_xref="PANTOM_DB:B230346M13"
REFERENCE	/db_xref="taxon:10090"
AUTHORS	/clone="B230346M13"

CDS

```

/sex="male"
/issue_type="corpora quadrigemina"
/clone_id="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
447..1952
/notes="unnamed protein product; beta-site APP cleaving
enzyme (MGD MGI:11346542, GI:NM_011792, evidence: BLASTN,
98%, match=3874)
putative"
/codon_start=1
/db_xref="GI:26337869"
/translation="MAPKAPK1LWVSGMLPAQCTHGLRPLPSGLAPPLGLRP
REDAESSEPRGRSFEVNDLRKSGGYVETVGSPEQTLNLDVSSNRAVG
AAPHLRHYRQSLSTYRDLKGVYVYQKMEGLGTLVSPHNPVTRVANI
AAITSDKFFPNSMWEGLGLVAIAIPDLSLEPDLVKOTHPNIPFLOLCA
GPNLOTBALAVSGSMIGIGIDHSLYTRLWYTPRIRMEYEVILVRIEIGDLM
DKENYVKSIVDSCTYNNLRPKVPEAAVKSIIKASSRKEPPDGFWMLEOLVMOAG
TTPWNI FVITSLTLMGEVTNOSFRITIILOQTLRPEDVATSDQDCYKRAVOSSTGT
VNGAVIMEGFYVFPDARRRIGFVASACHVDEPFAVAEFPVADMEDCGTNIPQT
DESTIMTIYVMAALCALFMLPLCLMVCMRCRLRHQDDFADISILK"

```

ORIGIN

Alignment Scores:  
 Pred. No.: 1,296-238  
 Score: 2211.00  
 Percent Similarity: 97.46%  
 Best Local Similarity: 96.30%  
 Query Match: 96.63%

Length: 4101  
 Matches: 417  
 Conservative: 5  
 Mismatches: 11  
 Indels: 0  
 Gaps: 0

US-10-726-967a-84 (1-433) x AK046175 (1-4101)

```

1 ThrGlnHisGlyIleArgLeuProLeuArgSerGlyLeuGlyValAlaProLeuGlyLeu 20
510 ACCCATCTGGCATCCGGCTCCCTTCGACGGCGCTGGCAGGCCACCCCTGGGCGT 569
21 GlnIleAsnLeuGluThrAspGluGluProGluGluProGlyValArgGlySerPheVal 40
570 AGGCTGCCCGGAGACCGACGAGAAATCGAGAGACCTGGCCGAGAGGCACTTTGAT 629
41 GlnMetValAspAsnLeuAlaGlyIleSerGlyGlnGlyTyrTyrValGlnMetCysVal 60
630 GAGATGATGACAACTCGAGGAGAAAGTCCGCGCAGGCTACTATGTGGAGATGACCGTA 669
61 GlySerProGluGlnThrIleAsnIleLeuValAspThrGlySerSerAsnPheAlaVal 80
690 GGCAGCCCCCAACAGCTCAACATCTGCTGACACGGGCACTAGTAACTTTGCAGTG 749
81 GlyAlaAlaProHisProPheLeuHisArgTyrTyrGlnArgGlnLeuSerSerThrTyr 100
750 GGGGCTGGCCCAACCTTTCTTCGATCGCTACTACACAGGAGGAGCTGTCACACATAT 809
101 ArgAspLeuArgGlyValTyrValProTyrThrGlnGlyIleTyrTrpGluGluLeu 120
810 CGAGACCTCCGAAAGGCTGTGTATGTCCCTTACACCGGAGCAAGTGGAGGGGAACTG 869
121 GlyThrAspLeuValSerIleProHisGlyProAsnValThrValArgAlaAsnIleAla 140
870 GGCACCGACCTGGTGGACATCTCTCATGCCCCCAACCTCTGCTGCTGCCAATTCCT 929
141 AlaIleThrGlnSerAspLeuPhePheHisAsnGlySerAsnTrpGluGlyIleLeuGly 160
930 GCCATCATCTGATCGACAGATTCTTCTCATATGTTCCAACTGGAGGGGCACTCTTGG 989
161 LeuAlaTyrAlaGluIleAlaArgProAspAspSerLeuGluProPheAspSerLeu 180
990 CTGGCCATATCTGATCGACAGGCTCGACGACCTCTTGGAGCCCTTCTTGCCTCTG 1049
181 ValIleGlnThrHisValProAsnLeuPheSerLeuGlnLeuGlyValAlaGlyPhePro 200
1050 GTGAAGCAACCACTCCCAACATCTTTTCTCTGACGCTCTGTGGCGCTGCTTCC 1109

```

```

QY 201 LeuAnginSerGluValIleAlaSerValGlyIleSerMetIleIleGlyIleAsp 220
DB 1110 CTCACACGACCGACGACCTGGCTCCGTTGGAGGAGCAATGATTCATTCATTCAC 1169
QY 221 HisSerLeuTyrThrGlySerLeuTrpTyrThrProIleArgGluTrpTyrTyrGlu 240
DB 1170 CACTGCTATACACGGGACAGCTCTGTATACACACCATCCGCGGAGTGGATATATGAA 1229
QY 241 ValIleIleValArgValGluIleAsnGlyIleAspLeuValMetAspCysArgGluTyr 260
DB 1230 GGATCATATGATCGTGGAAATCAATGATCAAGATCTCAAGATGACTGCAAGAGTAC 1289
QY 261 AsnTyrAspLeuSerIleValAspSerGlyThrThrAsnLeuArgLeuProIleVal 280
DB 1290 AACTACGACAGACATTTGTGACAGTGGACACCAACCTTCCTTCCCAAGAAAGTA 1349
QY 281 PheGluAlaAlaValIleSerIleValAlaSerSerThrGlnIlePheProAspGly 300
DB 1350 TTTGAAGCTGCCGTCAAGTCATCAAGGACAGCTCTCGACGAGAAAGTTCCCGATG 1409
QY 301 PheTrpLeuGlyGluGlnLeuValCysTrpGlnAlaGlyThrThrProTrpAsnIlePhe 320
DB 1410 TTTTGGCTTAGGGACAGCTGGTGTGCTGGCAAGGACGACGACCTTGGAACTTTTC 1469
QY 321 ProValIleSerLeuTyrIleMetGlyGluValThrAsnGlnSerPheArgIleThrIle 340
DB 1470 CCAGTCATTTTCACTTACTCATGGGTGAAGTACCAATCACTCTTCGATCAACATC 1529
QY 341 LeuProGlnGlnTyrIleAsnArgProValGluAspValAlaThrSerGlnAspAspCysTyr 360
DB 1530 CTTCCTGACATATACCTACGCGCGGTGAGGACGTGGCCACGTCACAGACATCTTAT 1589
QY 361 LysPheAlaIleSerGlnSerSerThrGlyThrValMetGlyValAlaIleMetGlyGly 380
DB 1590 AAGTTCGCTGTCTCAAGTCAATCAACGCGCAGTGTATAGGAGCCGTCTATCGAAGGT 1649
QY 381 PheTyrValValPheAspArgAlaArgLysArgIleGlyPheAlaValSerAlaCysHis 400
DB 1650 TTCTATGCTGCTTCGATTCAGACCCGAAAGCAATTTGGCTTGTCTGACGCTTCCAT 1709
QY 401 ValHisAspGluPheArgThrAlaAlaValGluGlyProPheValThrLeuAspMetGlu 420
DB 1710 GTGACGATGATTCATTCAGACGCGGACAGTGAAGGTCGTTGTATACGACAGATGAA 1769
QY 421 AspCysGlyTyrAsnIleProGlnThrAspGluSerThr 443
DB 1770 GACTGTGCTACACATTCCTCCAGACAGATGATCAACA 1808

RESULT 9
AK080498 3880 bp mRNA linear HTC 03-APR-2004
LOCUS AK080498.1 GI:26099278
DEFINITION Mus musculus 7 days neonate cerebellum cDNA, RIKEN full-length
enriched library, clone:A730059X08 product:beta-site APP cleaving
enzyme, full insert sequence.
ACCESSION AK080498
VERSION AK080498.1
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLES High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Kondo, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLES Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)

```

MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
20499374  
11042159  
3  
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,  
Komuro,H., Akiyama,J., Nishi,K., Kitamura,T., Tashiro,H., Itoh,M.,  
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,  
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,  
Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watabiki,M.,  
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsumura,S., Kawai,J.,  
Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A. and Hayashizaki,Y.,  
RIKEN integrated sequence analysis (RISA) system-384-Format  
sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
4  
The RIKEN Genome Exploration Research Group Phase II Team and the  
FANTOM Consortium.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409, 685-690 (2001)

JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
5  
The FANTOM Consortium and the RIKEN Genome Exploration Research  
Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation  
of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

JOURNAL  
AUTHORS  
6 (bases 1 to 3880)  
Adachi,T., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,  
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,  
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,  
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,  
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Komuro,H., Kouda,M.,  
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,  
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohnato,N.,  
Okazaki,Y., Saito,R., Saitoh,H., Sekai,C., Sekai,K., Sakazume,N.,  
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,  
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,  
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,  
Muramatsu,M. and Hayashizaki,Y.  
Direct Submission  
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
Kanagawa 230-0045, Japan (E-mail:genome-res@gsr.riken.jp,  
URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,  
Fax:81-45-503-9216)  
cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.  
Please visit our web site for further details.  
URL:http://genome.gsc.riken.jp/  
URL:http://fantom.gsc.riken.jp/  
Location/Qualifiers  
1..3880  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM\_DB:A730059K08"  
/db\_xref="taxon:10090"  
/clone="A730059K08"  
/tissue\_type="cerebellum"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="7 days neonate"  
450..1957  
/note="beta-site APP cleaving enzyme (MGI:MGI:1346542,  
GB|NM\_011792, evidence: BLASTN, 98%, match=3874)  
putative"

ORIGIN  
Alignment Scores:

Pred. No.: 3,726-236 Length: 3880  
Score: 2189.00 Matches: 417  
Percent Similarity: 97.46% Conservative: 5  
Best Local Similarity: 96.30% Mismatches: 11  
Query Match: 95.67% Indels: 1  
DB: 3 Gaps: 0

US-10-726-967a-84 (1-433) x AK080498 (1-3880)

QY 1 ThrGlnHisGlyIleArgLeuProLeuArgSerGlyLeuGlyAlaProLeuGlyLeu 20  
Db 516 ACCCATCTGGGATCCGGCTCCCTT-CGACGGGCTGGAGGCCACCCCTGGGCTG 574

QY 21 GlnIleAsnLeuGlnThrAspGlnGluProGlnGluProGlnGlyArgArgGlySerPheVal 40  
Db 575 AGGCTGCCCGGAGACCGACGAGGAATCGGAGGAGCGCTGGCCGAGAGCGACCTTGTG 634

QY 41 GlnMetValAspAsnLeuArgGlyLysSerGlyGlnGlyTyrTyrValGlnMetThrVal 60  
Db 635 GAGATGTGGACAACTGAGGGGAAAGTCGGCCAGGAGCTATGTTGGAGATGACGTA 694

QY 61 GlySerProGlnThrLeuAsnIleLeuValAspThrGlySerSerAsnPheAlaVal 80  
Db 695 GGACAGCCCCCAGACGAGCTCAACATCTGGTGGACAGGGCAGTAACTTTGGCAGTG 754

QY 81 GlnAlaAlaProHisProPheLeuHisArgTyrTyrGlnArgGlnLeuSerSerThrTyr 100  
Db 755 GGGGCTGCCCGACACCTTTCTCGCATCGCTACACAGAGGCGAGCTGCAGACATAT 814

QY 101 ArgAspLeuArgLysGlyValTyrValProTyrThrGlnGlySerThrGlnGluLeu 120  
Db 815 CGAAGCTCCGAAGGGTGTGTATGTCTTACCAACCGAGGAGAGTGGAGGGGAACTG 874

QY 121 GlyThrAspLeuValSerIleProHisGlyProAsnValThrValArgAlaAsnIleAla 140  
Db 875 GGCAACCACTGGGAGCATCTCCATCGGCCCAACGCTCACTGGCTGCACCATTGCT 934

QY 141 AlallethrGlnSerAspLysPhePheIleAsnGlySerSerThrGlnGlyIleLeuGly 160  
Db 935 GCCATCACTGAATCGAACGATTTTCATCAATGGTTCACACGGGAGGCACTCAAGG 994

QY 161 LeuAlaTyrAlaGlnIleAlaArgProAspAspSerLeuGluProPhePheAspSerLeu 180  
Db 995 CTGGCCTATGCTAGATTGGCAGGCCGAGCATCTTTGAGGCCCTTTTGACTCCCTG 1054

QY 181 ValLysGlnThrHisValProAsnLeuPheSerLeuGlnLeuGlyAlaGlyPhePro 200  
Db 1055 GTGAGCAGACCCACATCTCCCAATCTTTCCCTCAGCTGTGGCGCTGCCCTCCCC 1114

QY 201 LeuAsnGlnSerGluValIleAlaSerValGlyGlySerMetIleIleGlyIleAsp 220  
Db 1115 CTCACCAACAGCCAGGACCTGGCTGGTGGGAGGAGCATGATCATTTGGTATCGAC 1174

QY 221 HisSerLeuTyrThrGlySerLeuTyrTyrThrProIleArgArgGlnTyrTyrGlu 240  
Db 1175 CACTCGCTATACAGGGCAGGTCTCTGTACACCCATCCGGGGAGTGTATTATGAA 1234

QY 241 ValIleIleValArgValGlnIleAsnGlyGlnAspLeuLysMetAspCysLysGluTyr 260  
Db 1235 GTATCATTTGTAGTGTGGAAATCAATGTGCTCAAGATCTCAAGATGACGTCAAGAGGTATC 1294

QY 261 AsnTyrAspLysSerIleValAspSerGlyThrThrAsnLeuArgLeuProLysVal 280  
Db 1295 AACTGACAGAGCATTTGTGACACTGGGACCAACCACTTGGCTTGGCCAAAGAAATGA 1354

QY 281 PheGlnAlaAlaValLysSerIleLysAlaAlaSerSerThrGlnLysPheProAspGly 300  
Db 1355 TTGAGAGCTGGCGTCAAGTCCATCAAGGAGCGCTTCGACCGAGAAAGTTCCGGATGGC 1414

QY 301 PheTyrPheGlyGlnGlnLeuValCysTyrGlnAlaGlyThrThrProTyrAsnIlePhe 320  
Db 1415 TTTTGGCTAGGGGAGAGCGTGTGTGCTGGCAAGCAGGACGACCCCTTGGAAATTTTC 1474

Qy	321	ProvallieserleuYrleuMeG1g1ValThrxnsgInserPheAsg1leThrIle	340
Db	1475	CCAGTCAATTACTTAACTTAACTGAGGAGATCAACATCAAGTCTTCGCCATACATC	1534
Qy	341	LeuProlngIntrYrleuArProValg1uApVala1ThrSerGlnAspAspCysTyr	360
Db	1535	CTTCTCAGCAATACCTACCGGCCGCTGGAGGACGTGGCCAGCTCCCAAGAGACGTTCAC	1594
Qy	361	LysPhealalieserGInserSerThrG1yThrValMetG1yAlaval1leMetG1uG1y	380
Db	1595	AAGTTCGCTGCTCCAGTCATCCACGGGCACTGTATGAGGACCGCTCATCATGAAAGT	1654
Qy	381	PheYrValValPheAspArgAlaArgYrArg1leG1yPhealaval1serAlaCysHis	400
Db	1655	TTCTATGCTGCTTCGATCGAGCCCGAAAGCGAATGTGCTTGTGCTGTCAGCGCTTGCCAT	1714
Qy	401	ValHisAspG1upPheArgThrAlaAlaValG1uG1yProPheVal1ThrLeuAspMetG1u	420
Db	1715	GTGCACGATGAGTTTCAGACCGGCGGCACTGGAAGTCCGTTTGTTCACGCGACATGGAA	1774
Qy	421	AspCysG1yTyrAsn1leProGlnThrAspG1uSerThr	433
Db	1775	GACTGTGGCTACCAACATTCGCCACAGACATGATCAACA	1813
RESULT 10			
AK082230			
LOCUS			
DEFINITION	AK082230	3805 bp	mRNA
VERSION	AK082230.1	GI:26100512	
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			

REFERENCE	AUTHORS	COMMENT
JOURNAL	of 60,770 full-length cDNAs	
REFERENCE	Nature 420, 563-573 (2002)	
AUTHORS	6 (bases 1 to 3805)	
	Adachi, U., Aizawa, K., Akinura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiwamori, K., Hirooka, T., Hirozane, T., Horii, P., Imcanti, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishii, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Saesaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sobabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akashira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.	
TITLE	Direct Submission	
JOURNAL	Submitted (16-APR-2002)	
	Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Sushiro-cho, Tsukuba, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)	
FEATURES	CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.	
SOURCE	Please visit our web site for further details.	
	URL: http://genome.gsc.riken.jp/	
	URL: http://fantom.genb.riken.jp/	
	Location/Qualifiers	
	1..3805	
	/organism="Mus musculus"	
	/mol_type="mRNA"	
	/strain="C57BL/6J"	
	/db_xref="PANTOM DB:C230026008"	
	/db_xref="taxon:10090"	
	/clone_id="C230026008"	
	/clone_type="cerebellum"	
	/clone_id="RIKEN full-length enriched mouse cDNA library /dev stage="0 day neonate"	
	448..1955	
	/note="beta-site APP cleaving enzyme (MGD MGI:1346542, GB NM_011792, evidence: BLASTN, 98%, match=3874)	
	putative"	
ORIGIN		
	misc_feature	
	448..1955	
	/note="beta-site APP cleaving enzyme (MGD MGI:1346542, GB NM_011792, evidence: BLASTN, 98%, match=3874)	
	putative"	
Alignment Scores:	2,91e-235	3805
Pred. No.:	2181.00	417
Score:	97.01%	Conservative: 5
Percent Similarity:	95.86%	Mismatches: 11
Best Local Similarity:	95.32%	Indels: 2
Query Match:	3	Gaps: 0
DB:		
US-10-726-967A-84 (1-433) x AK082230 (1-3805)		
QY	1	ThrglnHlsGyltAlarGlupProleuYrSerGlyLeuGlyValAlaProleuGlyLeu 20
DB	511	ACCCATCTCGGATCGGCTGACCTTGACACGGGCTTGACGAGGCCACCTCGGCTG 570
QY	21	GuilAeAmLeuGluThrAspGluGluProGluGluProGlyYrGhrGlySerPheVal 40
DB	571	AGCGTGGCCCGGAGAACGACGAGAAATCGAGGAGCGCTGGCGGAGAGCGACTTGTG 630
QY	41	GlueVetAlaPAsnLeuArgGlyYrSerGlyGlnGlyYrYrValGluMetThrVal 60
DB	631	GAGATGTGTGACAACTGACGGGAAAGTCCGGCCACGGGCTACTATCTGAGAGTACCGTA 690
QY	61	GlySerProProGluThrLeuAmnLeuValAlaPThrGlyYrSerAsnPheAlaVal 80
DB	691	GGCAGCCCCCAAGACGCTCAACATCTCTGTGTGACACGGGACGATGTAATCTTGACATG 750

Qy 81 GlyAlaIaIaPProHISProPheLeuHISArgTYrTyrGlnArgGlnLeuSerSerThrTYr 100  
 Db 751 GGGGCTCCCCACACCCCTTTCCTGCAATCGCTACACAGAGGCGCTGTCACACATAT 810  
 Qy 101 ArgAspLeuArgIsgIyValTYrValProTYrThrGln-GlyLeuTrpGlnGluGlu 120  
 Db 811 CGAGACTCCGAAAGGGTGTGTATGTCTCCTACACCCAGGGGCAAGTGGAGGGAACT 870  
 Qy 120 UGIYThAspLeuValSerIleProHISGlyProAsnValThrValArgAlaAsnIleAl 140  
 Db 871 GGGGACCGACCTGTGTAGCATCTCCTCATGGCCCCAAGCATGCTGGGTCCAACTTGG 930  
 Qy 140 aAlaIleThGlySerAspLysPhePheIleAsnGlySerAsnTrpGlnGlyIleLeuG 160  
 Db 931 TGCCATCATGTAAATCGACAAAGTCTTCATCAATAGTTCACACTGGAGGGCATCTTAA 990  
 Qy 160 YLeuAlaTYrAlaGluIleAlaArgProAspAspSerLeuGluProPhePheAspSer 180  
 Db 991 GCTGGCTATGTGAGATTGCCAGGCCGACGACTCTTGGAGCCCTTCTTGCATCTCCCT 1050  
 Qy 180 UValIysGlnThHISValProAsnLeuPheSerLeuGlnLeuCyGlyAlaGlyPhePr 200  
 Db 1051 GGTGAAGCAGACCCACATTCCTCCAAACATCTTTCCTGCACTGTGGCGCTGGCTTCCC 1110  
 Qy 200 oLeuAsnGlnSerGlyUValLeuAlaSerValGlyGlySerMetIleIleGlyIleAs 220  
 Db 1111 CTTCAACACAGACCGAGGCACTGGCTCGTGGAGGGAGCATATCATTTGGTGTATCA 1170  
 Qy 220 PHisSerLeuTYrThGlySerLeuTrpTYrThrProIleArgArgGluTrpTYrTyrG 240  
 Db 1171 CCCTCGCTATACACGGGCGAGTCTCTGTGTACACACCATCCGGCGGAGTGGATATGA 1230  
 Qy 240 UValIleIleValArgValGluIleAsnGlyGlnAspLeuLysMetAspCyLeuGlyL 260  
 Db 1231 AGTGATCATTTGTACGTGTGAAATTCATGTCTCAAGATCTCAAGATGATCGAAGAG 1290  
 Qy 260 rAsnTYrAspLysSerIleValAspSerGlyThrThrAsnLeuArgLeuProLysIle 280  
 Db 1291 CAACCTACGACAGAGAGCATTTGTGACAGTGGACACCAACCTTCGTTGCCAAGAACT 1350  
 Qy 280 lPheGluAlaAlaValLysSerIleLysAlaAlaSerSerThrGluLysPheProAsp 300  
 Db 1351 ATTTGAAGTCCGCTCAAGTCCATCAAGGACAGCTCTCTCGAGGAGATTCCCGATG 1410  
 Qy 300 yPheTrp-LeuGlyGlnGlnLeuValCysTrpGlnAlaGlyThrThrProTPAsnIle 320  
 Db 1411 CTTTGGGCTTAGGGAGAGAGCTGTGTCTGCTGGACAGGAGCAAGCCCTTGGAACTTT 1470  
 Qy 320 hProValIleSerLeuTYrLeuMetGlyGluValThrAsnGlnSerPheArgIleThr 340  
 Db 1471 TCCCAAGTCAATTCATCTTACTCATGTGGTGAAGTCAACATCATGCTTCCGATACA 1530  
 Qy 340 lAlaLeuProGlnGlnTYrLeuArgProValGluAspValAlaThrSerGlnAspAsp 360  
 Db 1531 TCTTCTCTAGCAAAATCCTACGCGGGTGAAGAGAGTGGCCAGTCCCAAGACGACTTT 1590  
 Qy 360 yLysPhePheAlaIleSerGlnSerSerThrGlyThrValMetGlyValAlaIleMetG 380  
 Db 1591 ACAAGTGGCTGTCTCACAGTATCCACGGGACGCTTATGGGACCCCTCATCATGAG 1650  
 Qy 380 lYPheTYrValValPheAspArgAlaArgLysArgIleGlyPheAlaValSerAlaCy 400  
 Db 1651 GTTTCTATGTCTGTCTTCATTCAGGCCGCAAGCAATTTGGCTTGTCTGACGCTTCC 1710  
 Qy 400 lValAlaAspGluPheArgThrAlaAlaValGlnGlyProPheValThrLeuAspMet 420  
 Db 1711 ATGTGACGATAGTTCAGAGAGCGGCGAGTGGAAAGTCCGTTGTATACGCGACAGCAT 1770  
 Qy 420 lAspCyGlyTYrAsnIleProGlnThrAspGluSerThr 433  
 Db 1771 AAGACTGTGGTACAAACATTTCCCGACAGATGATGTAACA 1811

# RESULT 11 AK082317 LOCUS DEFINITION

AK082317 4048 bp mRNA linear HTC 03-APR-2004  
 Mus musculus 0 day neonate cerebellum cDNA, RIKEN full-length  
 enriched library, clone: C230037E16 product: beta-site APP cleaving  
 enzyme, full insert sequence.

## ACCESSION VERSION KEYWORDS SOURCE ORGANISM

AK082317.1 GI:26349644  
 HTC; CAP trapper.  
 Mus musculus (house mouse)  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

## REFERENCE AUTHORS JOURNAL MEDLINE PUBMED REFERENCE AUTHORS

Carninci, P. and Hayashizaki, Y.  
 High-efficiency full-length cDNA cloning  
 Meth. Enzymol. 303, 19-44 (1999)  
 99279253  
 10349636  
 2  
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
 Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new genes  
 Genome Res. 10 (10), 1617-1630 (2000)  
 20499374  
 11042159

## JOURNAL MEDLINE PUBMED REFERENCE AUTHORS

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,  
 Kono, H., Akiyama, J., Nishi, K., Kiteunai, T., Tashiro, H., Itoh, M.,  
 Sumi, N., Ishii, Y., Nakamura, S., Hayama, M., Nishino, T., Harada, A.,  
 Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashinaga, K.,  
 Fujiwake, S., Inoue, K., Togawa, Y., Izawa, K., Tanaka, T., Kawai, J.,  
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunura, S., Kawai, J.,  
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
 RIKEN integrated sequence analysis (RISA) system-384-format  
 sequencing pipeline with 384 multicapillary sequencer  
 Genome Res. 10 (11), 1757-1771 (2000)  
 20530913  
 11076661

## JOURNAL MEDLINE PUBMED REFERENCE AUTHORS

4  
 The RIKEN Genome Exploration Research Group Phase II Team and the  
 FANTOM Consortium.  
 Functional annotation of a full-length mouse cDNA collection  
 Nature 409, 685-690 (2001)

## JOURNAL MEDLINE PUBMED REFERENCE AUTHORS

5  
 The FANTOM Consortium and the RIKEN Genome Exploration Research  
 Group Phase I & II Team.  
 Analysis of the mouse transcriptome based on functional annotation  
 of 60,770 full-length cDNAs  
 Nature 420, 563-573 (2002)

## JOURNAL MEDLINE PUBMED REFERENCE AUTHORS

6 (bases 1 to 4048)  
 Adachi, U., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,  
 Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, M.,  
 Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T.,  
 Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,  
 Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Konda, M.,  
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,  
 Nakamura, M., Nishi, K., Nomura, K., Nunazaki, R., Ohno, M., Ono, N.,  
 Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,  
 Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shizaki, T.,  
 Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akihara, S.,  
 Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,  
 Muramatsu, M. and Hayashizaki, Y.

## TITLE JOURNAL

Direct Submission  
 Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of  
 Physical and Chemical Research (RIKEN), Laboratory for Genome  
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
 RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
 Kanagawa 230-0045, Japan (E-mail: genome-resgsc.riken.jp,  
 URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,  
 Fax: 81-45-503-9216)  
 CDNA library was prepared and sequenced in Mouse Genome

## COMMENT

Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.

Please visit our web site for further details.  
URL: <http://genome.gsc.riken.jp/>  
URL: <http://fantom.gsc.riken.jp/>  
Location/Qualifiers

## FEATURES

## SOURCE

1..4048

/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM\_DB:C230037E16"  
/db\_xref="taxon:10090"  
/clone="C230037E16"  
/tissue\_type="cerebellum"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="0 day neonate"  
451..1854  
/note="unnamed protein product; beta-site APP cleaving  
enzyme (MGD|MG1:1346542, GB|NM\_011792, evidence: BLASTN,  
98%, match=3874)  
putative"  
/codon\_start=1  
/protein\_id="BAC38462.1"  
/db\_xref="GI:26349645"  
/translation="MAPALHMLLWVSGMLPAOGTHLIRLPLRSLGAPPLGLRLP  
RTDESEPERGRGSFVEMVDNLRGKSGGYVEMTVGSPQTLLIIVDTGSSNFAVG  
AAPHPFIRHYOROLSTYRDLRKGVVYVYQKMGELGDTLIPHPNVTYANI  
AAITSDKRFINGSMWEGILGIAVAIAPDLSLPPRSLVKOTHIPIPSLOLGA  
GFLNQTALASVSGSMIIGIDHSLYTSLSLWTTPIRREMYEVITVREINGDLKM  
DKETKEKPEDFMLGEOVLVQWAGTTPMNIIPVISLYLMGVTNDSFRITLIPQDYLK  
PVEDVATSGDDCYKPAVSSQSTYGTWGAIVMEGFYVVDRAKRIGFVSAHVDEF  
RTAAVEGPTVADMECGYNIPTDESLMTIAYMAAICALFMLPLCLMVCQMRCLR  
CLRHQHDHPADDISILK"

## CDS

/note="unnamed protein product; beta-site APP cleaving  
enzyme (MGD|MG1:1346542, GB|NM\_011792, evidence: BLASTN,  
98%, match=3874)  
putative"  
/codon\_start=1  
/protein\_id="BAC38462.1"  
/db\_xref="GI:26349645"  
/translation="MAPALHMLLWVSGMLPAOGTHLIRLPLRSLGAPPLGLRLP  
RTDESEPERGRGSFVEMVDNLRGKSGGYVEMTVGSPQTLLIIVDTGSSNFAVG  
AAPHPFIRHYOROLSTYRDLRKGVVYVYQKMGELGDTLIPHPNVTYANI  
AAITSDKRFINGSMWEGILGIAVAIAPDLSLPPRSLVKOTHIPIPSLOLGA  
GFLNQTALASVSGSMIIGIDHSLYTSLSLWTTPIRREMYEVITVREINGDLKM  
DKETKEKPEDFMLGEOVLVQWAGTTPMNIIPVISLYLMGVTNDSFRITLIPQDYLK  
PVEDVATSGDDCYKPAVSSQSTYGTWGAIVMEGFYVVDRAKRIGFVSAHVDEF  
RTAAVEGPTVADMECGYNIPTDESLMTIAYMAAICALFMLPLCLMVCQMRCLR  
CLRHQHDHPADDISILK"

## ORIGIN

## Alignment Scores:

Pred. No.: 2,69e-217 Length: 4048  
Score: 2023.00 Matches: 384  
Percent Similarity: 89.84% Conservative: 5  
Best Local Similarity: 88.68% Mismatches: 10  
Query Match: 88.42% Indels: 34  
DB: 3 Gaps: 1

US-10-726-967a-84 (1-433) x AK082317 (1-4048)

QY 1 THRGINHIIGLYIYIARGLIUPROLEUARGSERGLYLEUGLYIYALAPROLEUGLYLEU 20  
DB 514 ACCCATCTGGCATCTCGACTGCCCCCTTCCAGCGGCTGGCAGGCGCACCCCTGGGCTG 573  
QY 21 GIUILEASLEUGLUTHRASPGLUGLUPROGLUGLUPROGLYARGLRGISERPEHA 40  
DB 574 AGGCTCCCCGGGAGACCAACAGAGAAATCGAGGAGCTGGCCGGAGAGGACGCTTTGG 633  
QY 41 GIUMETVALASPALEUARGLYLYSERGLYGLINGLYTYTYRYVALGIUMETTHRVA 60  
DB 634 GAGATGCTGGACCACTGAGGGGAAAGTCCGGCCAGGGCTACTATGTGAGATGACCGTA 693  
QY 61 GLYSERPROGLINTHIRLEUANIILEUVALASPTHGLYSERERANPEHAVAL 80  
DB 694 GGCAGGCCCCCAAGACGCTCAACATCTGTGTGAGACCGGCGAGTGAACCTTTGACATG 753  
QY 81 GLVALAIALPROHISPROPHILEUHIISARGTYTYTYGLINARGLINLEUSERSERTHTYR 100  
DB 754 GGGGCTGGCCCAACCTTTCTGCACTGCTACTACAGAGGACGCTGCTCAGACATAT 813  
QY 101 ARGASPLEUARGLYGLYVALYTYRYVALYTYRYTHRGINGLYLYYRTPGUGLUGLEU 120  
DB 814 CGAGACTCCGAAAGGCTGTGTATGTGCTTACACACGAGGCAAGTGGAGGGGAACTG 873  
QY 121 GLYTHASPLEUVALSERILEPROHISGLYPROHISVALTHRYVALARGAIAANILEA 140

DB 874 GGACCCGACCTGTGAGACATCCCTCATAGGCCCAACGTCATCTGCTGCCAATCTGCT 933  
QY 141 AAIETHRGLUSERAPLYSPHEPHELEANSGLYSERANTPGULYILEUGLY 160  
DB 934 GCCATCACTGATGAGCAAGATTCTTCAATAGTTCACCTGGAGGAGCATCTTAGGG 993  
QY 161 LEUAIYRALAGLUILEAIALARGPROASPAPSERLEUGLUPROPHASPSERLEU 180  
DB 994 CTGGCTTATGCTGATGATGCCAGGCCCAACATCTTTGGAGCCCTTCTTGACCTCCG 1053  
QY 181 VALLYSGINTHRHISVALPROASNLEUPHESERLEUGLNUCYGVALIAGLYPHEPRO 200  
DB 1054 GTGAAGCAGACCCACATCTCCCAACATCTTTCCCTGACGCTGTGGGCGCTGCTCCCC 1113  
QY 201 LEUANGINSERGLUVALLEUVALASERVALGYGLYSERMETILEGLYIYILEASP 220  
DB 1114 CTCAACCGAGCCGAGGCGCTGGCTGGGAGGAGCAATGATGATGGTGGATTCAC 1173  
QY 221 HISSELEUARYTHRGISERLEUTRPTYTRHPROLIETARGATGLUTRPTYTYRGLU 240  
DB 1174 CACTGCTATACACGGGAGCTCTCTGTACACACCCATCCGGGAGTGTGATTTAGAA 1233  
QY 241 VALILEILEVALARGVALGLUILEANGLYINAPLEUARYMERASPVALYSGLUTYR 260  
DB 1234 GTGATCATTTGATCGTGTGAATCATATGTCATCAAGATTCATCAAGATGACCTGCAAGAG- 1290  
QY 261 ASNTYRASPILYSERILEVALASPSERGLYTHRTHASLEUARGLEUPROLYSILYVAL 280  
DB 1290 ----- 1290  
QY 281 PHEGLUIAIALAVALLYSERILEYSAIALASERSETHRGULYSPHEPROASPGLY 300  
DB 1291 -----ACGAGAAATGCCGATGGC 1311  
QY 301 PHETRPLEUGLYGLUINLEUVALCYSTRPGINALAGLYTHRTHPROTRPASNILEPHE 320  
DB 1312 TTTTGCTATAGGGAACAGCTGGTGTGTGGCAGACGACGACGACCCCTTGGAACATTTTC 1371  
QY 321 PROVALIISERLEUTRYLEUMETGLYGLUVALTHRAANGINSERPHETARGILETHRILE 340  
DB 1372 CCAGTCATTTCCTTACTCTCATGTGGTGAAGACCAATCACTCTTCGCAATCCATC 1431  
QY 341 LEUPROGLINGINTYRLEUARGPROVALIGUASPVALAATHRSEGLNAPASPCYSTRY 360  
DB 1432 CTTCCTCAGCAATACCTACGCGGAGGAGACGATGGCCAGCTCCCAAGACGACTGTTC 1491  
QY 361 LYSPEHAALIESERGLNSERSETHRGLYTHRVALMETGLYALVALILEMETGLUGLY 380  
DB 1492 AAGTTCGCTGTCTCAGATCCACGAGGCACTGTATTGGAGCGGTATCATGGAAGGT 1551  
QY 381 PHEYRYVALVALPHEASPARGLAARGLYSARGILEGLYPHEALVALISERLAACYRHIS 400  
DB 1552 TTCATATGCTGCTGTGATGACGAGCCGAAAGCAATGTGCTTTGCTGACGCGCTTGCA 1611  
QY 401 VALHISAPRGULUPHARGTHRVALAVALAGLUGLYPROPHETVALTHIRLEUASMETGLU 420  
DB 1612 GTGCGAGATGATTCAGAGCGGCGAGGAGGATCGTTTGTTCACGACATGGA 1671  
QY 421 AEPCYSGLYTYRASNILEPROGLINTHRASPGIUSERTHR 433  
DB 1672 GACTGTGCTACACATTTCCCAAGACAGATGATGATCA 1710

RESULT 12  
AK079814 3147 bp mRNA linear HTC 03-APR-2004  
LOCUS AK079814 Mus musculus 0 day neonate thymus cDNA, RIKEN full-length enriched  
DEFINITION library, clone:AA30076807 product:beta-site APP cleaving enzyme,  
full insert sequence.  
ACCESSION AK079814  
VERSION AK079814.1 GI:26098773  
KEYWORDS HTC, CAP trapper.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus



REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS

1  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)  
99279253  
10349636

2  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
20493374  
110429159

3  
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Taahito, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Wachihi, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
20530913  
11076861

4  
The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409, 685-690 (2001)  
5

5  
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)  
6 (bases 1 to 3147)

JOURNAL  
REFERENCE  
AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Nunazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sugabe, Y., Tagami, M., Tagawa, A., Takenashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.  
Direct Submission

TITLE  
JOURNAL

Submitted (16-Apr-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Sueni-ro-cho, Tsunumi-ku, Yokohama, Kanagawa 230-0045, Japan [E-mail: genome-research.riken.jp, fax:81-45-503-9216]  
URL: http://genome.gsc.riken.jp/, Tel:81-45-503-9222,

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
Please visit our web site for further details.  
URL: http://genome.gsc.riken.jp/  
URL: http://fantom.gsc.riken.jp/.  
Location/Qualifiers  
1. 3147  
/organism="Mus musculus"  
/mol\_type="mRNA"

FEATURES  
SOURCE

	misc_feature	
	1. 3147	
	/note="beta-site APP cleaving enzyme (MGD) MG1:1346542,	
	GB NM_011792, evidence: BLASTN, 98%, match=3874"	
ORIGIN		
		/sestrain="CS7BL/6J"
		/db_xref="PANTOM,DB:A430076B07"
		/db_xref="taxon:10090"
		/clone="A430076B07"
		/tissue_type="thymus"
		/clone_lib="RIKEN full-length enriched mouse cDNA library
		/dev_stage="0 day neonate"
Alignment Scores:		
Pred. No.:	2,45e-162	3147
Score:	1537.00	Matches: 302
Percent Similarity:	77.97%	Conservative: 6
Best Local Similarity:	76.46%	Matches: 12
Query Match:	67.18%	Gaps: 75
DB:	3	Indels: 2
US-10-726-967A-84 (1-433) x AKO79814 (1-3147)		
QY	51 GYGLNGLYTYTYRYVALGIWETHTVALGISERPROPGILNTHR-----	66
DB	86 GSAAGATTCACCAAGACACAGAGCCGGTGGGTCCAGAGCCGTCTGTGCATGTCTTT	141
QY	67 -----LEUANNILEUVALAAPTTHGISERSERANPHELA	79
DB	146 TCTCTCCCTTCTCTGCACAGCTCAACATCTGGTGAGACGGGCACTGTAATCTTCA	207
QY	80 VALGVALAALAPROHISPROPHLEUHSARGTYTYRGINARGINLEUSERETHR	99
DB	206 GTGGGGGCTGCCCAACACCTTCTCTGCATCTACTACCAAGGACAGCTGCCAGACA	261
QY	100 TYRARGAPLEUARGYSGYVALTYRYVALPROTYRTHRGINGLYLYETPGLUGLYGLU	111
DB	266 TATGAGAACTTCCGA-----	281
QY	120 LEUGLYTHRANPLEUVALSERILEPROHISGLYPROANVALTHRVALARGALASNNILE	131
DB	280 -----	281
QY	140 ALALALLETHRGUSERAPLYSPHEPHEILEANGLYSERANTPGLUGLYILEU	155
DB	280 -----	281
QY	160 GLYLEUVALTYRALAGLUILEALARG-PROASPAPSERLEUGIUPROPHAPSE	175
DB	281 -----AAGCCGACACACTTTGGAGCCCTTTTGACTC	311
QY	179 RLEUVALLYSGINTHHSIVALPROANPLEUHPHESERLEUGINLEUCYGLVALAGLYPH	195
DB	317 CCTGGTGAAGACAGACCACATTCACCAATCTTTTCTCTGCAGCTGTGGCGCTGCTT	376
QY	199 EPRLEUANGINSENGIUALLEUVALSERVALGLYSEKMETLEILEGLYGLYI	215
DB	377 CCCCCCTCAACCAAGAGGAGCACTGACCTGGTGGAGAGGACATGATCTGGTGGTAT	433
QY	219 EAPHSISERLEUYRTHGLYSERLEUTPTYRTHRPROILEARGARGIUTPTYRY	235
DB	437 CGACCACTGCTATACAGGGGCAAGCTCTGGTACACACCATCCGGCGGAGTGGTATTA	499
QY	239 RGLUVALLEILEVALARGVALGLUILEANGLYGINAPLEULYSMETAPCYLYSGI	255
DB	497 TGAAGGATCATTTGACGTGGGAATCAATGGTCAAGATGTCAAGATGACTCCAAGGA	555
QY	259 UTYRANTRYAPLYSERILEVALASPSEGLYTHRTHRANPLEUARGLEUPPOLYSLY	275
DB	557 GTRCAACTACGACAAAGACATTTGGACAGTGGGACCAACCACTTGCTTGGCCAAAGAA	616
QY	279 EVALPHEGLUALAVALVALYSERILEVALALASERETHRGUYSRPHPEPRAE	299
DB	617 AGTATTTGAAGGTGGCGTCAAGTCCATCAAGGACAGCTCTCGACGAGAAAGTTCCGGA	676



QY 299 pGlyPheTLeuGlyGluGlnLeuValCysTrpGlnAlaGlyThrThrProTrpAsn11 319  
 Db 677 TGGCTTTGGTAGGGGAGACAGTGGTGTGGTGGCAAGGACGACCCCTGGAAACAT 736  
 QY 319 ePheProVal11IleSerLeuTyrLeuMetC1yGluVal11TrpAngInSerPheArg11Leh 339  
 Db 737 TTTCCAGCATTTCACTTACCTCAATGGGTGAAGTCAACATCAATGCTTCCGATCAG 796  
 QY 339 rIleLeuPProGlnGlnTyrLeuArgProValGluAspVal1A1aThrSerGlnAspAspCy 359  
 Db 797 CATCTCTCTCCAGCANTACTTACGCGCGGTGAGAGACGTCGACCTCCCAAGACAGCAG 856  
 QY 359 sTyrTrpPheAla1IleSerGlnSerSerThrGlyThrValMetGlyAlaVal11LeuMetC1 379  
 Db 857 TTACAAGTTCGGTGTCTCAAGTCATCCAGGGCAGCTGTTATGGAGCGCTCATCATGGA 916  
 QY 379 uGlyPheTyrVal1ValPheAspArg1A1aArgLysArg11LeGlyPheAlaVal1Ser1aCy 399  
 Db 917 AGGTTCTATGTCCTCTGATCGAGCCGGAAGCCGAAATGGCTTTGCTGTACAGCCCTTG 976  
 QY 399 sH1sValH1sAspGluPheArgThrAla1aValGluGlyProPheValThrLeuAspK 419  
 Db 977 CCAATGTCAGATGATTCAGAGCGCGGAGCTGGAAGTCCGTTTGTATCGGACAGAT 1036  
 QY 419 tGluAspCyGlyTyrAsn11eProGlnThrAspGlnSerThr 433  
 Db 1037 GGAAGACTGTGGTACAACTTCCCAAGACATGATGATCAACA 1079  
 RESULT 13  
 BUS57282 928 bp mRNA linear EST 16-SEP-2002  
 LOCUS BUS57282  
 DEFINITION AGENCOURT 10253167 NIH MGC 109 Homo sapiens cDNA clone  
 IMAGE:6585124 5', mRNA sequence.  
 ACCESSION BUS57282  
 VERSION BUS57282.1 GI:22907578  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 928)  
 NIH-MGC http://mgi.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgabs@emil.nih.gov  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone Distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LCM2796 row: j column: 04  
 High quality sequence start: 28  
 High quality sequence stop: 688  
 Location/Qualifiers  
 1..928  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6585124"  
 /issue\_type="teratocarcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: ovary; Vector: pOTB7, Site\_1: EcoRI; Site\_2:  
 XhoI; CDNA made by oligo-dT priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GGCAAGGAGG). Library constructed by Ling Hong in the  
 Laboratory of Gerald M. Rubin (University of California,  
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
 Superscript II RT (Life Technologies). Note: this is a

ORIGIN NIH\_MGC Library."  
 Alignment Scores:  
 Pred. No.: 1,836-150 Length: 928  
 Score: 1425.00 Matches: 285  
 Percent Similarity: 95.99% Conservative: 2  
 Best Local Similarity: 95.32% Mismatches: 6  
 Query Match: 62.28% Indels: 6  
 DB: 5 Gaps: 2  
 US-10-726-967a-84 (1-433) x BUS57282 (1-928)  
 QY 131 ProAsnVal1ThrValArgAlaAsn11LeAla1aThrGlnSerAspLysPhePhe11e 150  
 Db 36 CCACAGCAGCAGTGTGGTGGCCAACTTCTGCATCACTGAATCAAGACAGTTCTTCAATC 95  
 QY 151 AsnGlySerAsnTrpG1yG1yleuGlyLeuAlaTyrAlaGlu11eAlaArgProAsp 170  
 Db 96 AACGGCTCCAACTGGGAAGGATCTCGGGGCTGGCTTATGATGATTCAGAGCTGAC 155  
 QY 171 AspSerLeuGluProPheAspSerLeuVal1yGlnThrH1sVal1ProAsnLeuPhe 190  
 Db 156 GACTCCCTGAGACCTTTCTTTCATCTCTGTGAAGACAGACCAGTTCCCAACCTTTC 215  
 QY 191 SerLeuGlnLeuCyGlyAla1aGlyPheProLeuAsnGlnSerGluVal1LeuAlaSerVal 210  
 Db 216 TCCCTGAGCTTTGAGTGTGGCTTCCCTCCCAACAGTGTGAAGTGTGGCTTGTGC 275  
 QY 211 GlyGlySerMet11e1eGlyGly11eAspH1sSerLeuTyrThrGlySerLeuTyr 230  
 Db 276 GAGGAGAGCATATCATTTGAGGTATTCAGACCTGCTGTACAGGAGAGTCTGTGAT 335  
 QY 231 ThrPro11eArgArgGluTyrTyrTyrGluVal11e11eValArgValGlu11eAsnGly 250  
 Db 336 ACACCATTCGGCGGAGGTGATATGAGTCATCATTTGCGGGTGGAGATCAATGGA 395  
 QY 251 GlnAspLeuLysMetAspCyAlaGlyGluTyrAsnTyrAspLysSer11eValAspSerGly 270  
 Db 396 CAGGATCTGAAATGAGTCAGAGAGATCAACTGTGACAAAGCATTTGTGACAGTGGC 455  
 QY 271 ThrThrAsnLeuArgLeuProLysVal1PheGluAla1aVal1ySer11eLysAla 290  
 Db 456 ACCACCACTTCTGTTTGGCCAAAGGTTTGAAGCTGAGCTGAATCCATCAAGGA 515  
 QY 291 AlaSerSerThrGluLysPheProAspGlyPheTyrLeuGlyGlnLeuValCyStrp 310  
 Db 516 GCCTCTCAGGAGAGAGTTCCTGATGTTCTGGCTAGAGAGAGAGCTGTGTGCTGG 575  
 QY 311 GlnAlaGlyThrThrProTrpAsn11ePheProVal11IleSerLeuTyrLeuMetGlyGlu 330  
 Db 576 CAAGCAGGACCAACCCCTTGGAAATTTCCAGATCTCACTTAACTTAATGGGTGAG 635  
 QY 331 ValThrAsnGlnSerPheArg11eThr11eLeuProGlnGlnTyrLeuArgProValGlu 350  
 Db 636 GTTACCAACAGCTCTTCCGATCAACATCTTCCGACATTACTCGGCGCAGTGA 695  
 QY 351 AspValAla1aThrSerGlnAspAspCyTyrTrpPheAla1IleSerGlnSerSerThrGly 370  
 Db 696 GATGTGGCAGATCCCAAGACGATGTTCAAGTTGGTCCATCTCAAGATCAACGAGGCG 755  
 QY 371 ThrValMetGlyAlaVal11eMetGlyGlyPheTyrVal1ValPheAspArg1a-Argly 390  
 Db 756 ACTGTTATGGAGCTGTTCATGAGGGCTTCACTGATGTTGATGGGCGCGGAA 815  
 QY 390 sArgGly-GlyPheAlaVal1-SerAlaCyH1sValH1sAspGluPheArgThrAla1a 409  
 Db 816 ACGAATGGAGCTTGTGACAGCGCTTCCATGTGACAGATGAATCAAGGACGCGAC 875  
 QY 410 ---ValGluGlyPro-PheValThrLeu---AspMetGluAspCyGly 423  
 Db 876 CGGTGAAGAGGCGCTTTTGTCACTTGAACATGGAAGACTGTGGG 924

RESULT 14  
 BUI28383 1001 bp mRNA linear EST 25-NOV-2002  
 LOCUS 603113984F1 CSECHL20 Gallus gallus cDNA clone CHER65g21 5', mRNA  
 DEFINITION sequence.  
 ACCESSION BUI28383  
 VERSION BUI28383.1 GI:25339782  
 KEYWORDS EST  
 SOURCE Gallus gallus (chicken)  
 ORGANISM Gallus gallus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Aves; Neognathae; Galliformes; Phasianidae;  
 Phasianinae; Gallus.  
 1 (bases 1 to 1001)  
 Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,  
 Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.  
 A Comprehensive Collection of Chicken CDNAs  
 Curr. Biol. 12 (22), 1965-1969 (2002)  
 22335534  
 MEDLINE 12445392  
 PUBMED  
 COMMENT Contact: Simon Hubbard  
 Department of Biomolecular Sciences  
 University of Manchester Institute of Science and Technology  
 (UMIST)  
 PO Box 88, Manchester, M60 1QD, UK  
 Tel: 01612008930  
 Fax: 01612360409  
 Email: Simon.Hubbard@umist.ac.uk.

FEATURES  
 source  
 1..1001  
 Location/Qualifiers  
 /organism="Gallus gallus"  
 /mol\_type="mRNA"  
 /strain="White Leghorn, Hixex"  
 /db\_xref="taxon:9031"  
 /clone="CHST65g21"  
 /dev\_stage="36"  
 /lab\_host="DH10B"  
 /clone\_1fb="CSECHL20"  
 /note="Organ: limb; Vector: pBluescript II KS(+); Site 1:  
 EcoRI, Site 2: NotI; Modification of pBluescript II KS(+)  
 (Stratagene) vector to accommodate cDNA produced with the  
 1-trimmed protocol (Construction of uni-directionally  
 cloned cDNA libraries from messenger RNA for improved 3'  
 end DNA sequencing by Glenn Fu, et al. U.S. Patent #  
 6,387,624). Cut pBluescript II KS(+) with NotI and EcoRI.  
 ligate in double stranded adaptor containing BamI and  
 BamH sites (5'gaccgcgcagcccgatccgaaaaaag)  
 (5'aattcttttctcgatccg99gctcgacgc)"

ALIGNMENT SCORES:  
 Pred. No.: 8 51e-145 Length: 1001  
 Score: 1375.50 Matches: 287  
 Percent Similarity: 89.61% Conservative: 15  
 Best Local Similarity: 85.16% Mismatches: 32  
 Query Match: 60.12% Indels: 8  
 DB: 5 Gaps: 0

US-10-726-967a-84 (1-433) x BUI28383 (1-1001)

QY 55 TyrValGluMetThrValGlySerProProGlnThrLeuAsnIleLeuValAspThrGly 74  
 Db 2 TACGTGAGATGACGCTGGGACGCCCCCGAGAGCTCAATATCTCGTGGACACTGGG 61  
 QY 75 SerSerAspPheAlaValGlyAlaProHisPhePheLeuHisArgTyrTyrGlnArg 94  
 Db 62 AGCACTAATTTTGGCTGGAGACTGACCTCACCTTTCTCCGAGATACTACACGCGG 121  
 QY 95 GlnLeuSerSerThrTyrArgAspLeuArgGlyValTyrValProTyrThrGlnGly 114  
 Db 122 CAGCTGTCCAGACACTTACCTGACCTGCGGAAAGGTGTGTATGTGCTTACACCA-GGC 180  
 QY 115 LysTrpGlnGlyGluLeuGlyThrAspLeuValSerIleProHisGlyProAsnValThr 134

Db 161 AAGTGGAGAGGAGGAGCTGGGACCTGATCCATCTCCCACTGGCCCCCAAGCTACT 240  
 QY 135 ValArgAlaAsnIleAlaIleThrGluSerAspPhePheIleAsnGlySerAsn 154  
 Db 241 GTCAAGCCCAACATCGCTGCCATCACGGAGTCACAAATTTCTTCACTCATGTGCTCCAC 300  
 QY 155 TrpGlnGlyIleLeuGlyLeuAlaTyrAlaGluIleAlaArgProAspAspSerLeuGlu 174  
 Db 301 TGGAGAGGATCTCGGGGCTGGCTTATGCTAGATCCACAGCGCTGCACAGACGCTGGAG 360  
 QY 175 ProPhePheAspSerLeuValIleValGlyThrHisValProAsnLeuPheSerLeuGlu 194  
 Db 361 CCTTTTGTATCTCCCTGGTGAACAGACCCAGGTGCCCAACATCTTCTCCCTCAGCTT 420  
 QY 195 CysGlyAlaGlyPheProLeuAsnGlnSerGlyValIleuAlaSerValGlyGlySerMet 214  
 Db 421 TGTGGGGCAGGCTTCTACACCCACAGACAGAGACTTGGCATCAGTGGAGGAGCATG 480  
 QY 215 IleIleGlyGlyIleAspHisSerLeuTyrThrGlySerLeuTyrThrProIleArg 234  
 Db 481 ATCATTTGTGTCATCGACCGCTCGTGTATGTGGTGACATCTGTGTACACACCATCCGG 540  
 QY 235 ArgGluTrpTyrTyrGluValIleIleValArgValGluIleAsnGlyIleAspLeuLys 254  
 Db 541 AAGAGTGTGATCTACGAGGTGATATGTCAAGCTGAGGTCAATGGGACGACCTGAAC 600  
 QY 255 MetAspCysIleGlyIleTyrAsnTyrAspLysSerIleValAlaSerGlyIleThrAsnLeu 274  
 Db 601 ATGACACTGCAAAAGTCAACTATGACACAGATATCGTGACAGTGGCACCAACCACTTC 660  
 QY 275 ArgLeuProIleLysValIleGluAlaAlaValIleSerIleLysValAlaIleSerThr 294  
 Db 661 AGGTGTCGGAAGAGGAGTGT-GAGGCGCAGTGAATAATCAATCAAAACAGCTTCTTGACA 719  
 QY 295 GluLysPheProAspGlyPheTrpLeuGlyGluGlnLeuValCysTrpGlnAlaGlyThr 314  
 Db 720 GAGAAAGTCCCAATAGGCTTCTGGCTGGGGAGACAGTGTGTTGCTGGACGTCGACCA 779  
 QY 315 ThrProTrpAsnIlePheProValIleSerLeuTyrLeuMetGlyGluValThrAsnGln 334  
 Db 780 C--CCCTGGGACATCTTCCCTGTTGTCTCCCTTACCTGATGGGGAGGCGCAC-AAACAG 836  
 QY 335 SerPheArgIleThrIleLeuProGlnGlnTyrLeuArgProValGluAspValAlaThr 354  
 Db 837 TCTTCCGATCCACATCTTCTCCACGAACTGCGCCAGTGAAGAACTGTGGCCACT 896  
 QY 355 SerGlnAspAsp-Cys-TyrLysPheAlaIleSerGlnSerSer-ThrGlyThrValMet 373  
 Db 897 CTCAGGATGACCTGCTTACAGGTTGACAACTCTCAGGCTCCACAGCAGCATGCTCATG 956  
 QY 374 GLyAlaValIleMetGluGlyPheTyrValValPheAspArgAla 388  
 Db 957 GGGTGTGTATCCAGAGGGGTCCACCTGGC-TTTGACCGGGGCC 1000

RESULT 15  
 BUI502697 950 bp mRNA linear EST 12-SEP-2002  
 LOCUS AGENCOURT\_10023278 NIH\_MGC\_40 Homo sapiens cDNA clone IMAGE:5481022  
 DEFINITION 5', mRNA sequence.  
 ACCESSION BUI502697  
 VERSION BUI502697.1 GI:22807818  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 950)  
 NIH-MGC http://mgi.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 JOURNAL  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs-r@mail.nih.gov

Tissue Procurement: DCTD/DTF  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
 Plate: LNCM2661 row: 1 column: 15  
 High quality sequence stop: 620.  
 Location/Qualifiers

# FEATURES

1..950  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6481022"  
 /tissue\_type="carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_40"  
 /note="Organ: prostate; Vector: pOTB7; Site\_1: XhoI;  
 Site\_2: EcoRI; cDNA made by oligo-dT priming.  
 directionally cloned into EcoRI/XhoI sites using the  
 following 5' adaptor: GGCACGAG(G). Library constructed by  
 King Hong in the laboratory of Gerald M. Rubin (University  
 of California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies).  
 Note: this is a NIH\_MGC Library."

# ORIGIN

Alignment Scores:  
 Pred. No.: 1,22e-143 Length: 950  
 Score: 1365.00 Matches: 277  
 Percent Similarity: 94.59% Conservative: 3  
 Best Local Similarity: 93.58% Mismatches: 10  
 Query Match: 59.66% Indels: 7  
 DB: 5 Gaps: 2

US-10-726-967a-84 (1-433) x BU502697 (1-950)

QY 58 MetThrValAlaGlySerProPheGlnThrLeuAsnIleLeuValAspThrGlySerSerAsn 77  
 DB 1 ATGACCGTGGGAGAGCCCGGAGAGCTCAACATCTGTGGATACAGGAGAGCACTAAC 60  
 QY 78 PheAlaValAlaAlaProHisProPheLeuHisArgTyrTyrGlnArgGlnLeuSer 97  
 DB 61 TTGGAGTGGGTGCTGCCCCCCTCTGCTGCTACTACGACGAGGAGCTGTGCC 120  
 QY 98 SerThrTyrArgAspLeuArgLysGlyValTyrValProTyrThrGlnGlyLysTrpGlu 117  
 DB 121 AACCACTACCGGAGAGCTCCGAGAGGTGTGTATGTCCCTACACCCAGGAGCAAGTGGAA 180  
 QY 118 GlyLeuLeuGlyThrAspLeuValSerIleProHisGlyProAsnValThrValArgAla 137  
 DB 181 GGGAGAGCTGGGAGAGCCGAGCTGTAGCAATCCCATGAGCCCACTGAGCTGTGCC 240  
 QY 138 AsnIleAlaAlaIleThrGlnSerAspLysPhePheIleAsnGlySerAsnTrpGluGly 157  
 DB 241 AACATTGCTGGCCATCACTGAATCAGACAAAGTTCTTCATCAACGGCTCCAACTGGAAAGGC 300  
 QY 158 IleLeuGlyLeuAlaTyrAlaGluIleAlaArgProAspAspSerLeuGluProPhePhe 177  
 DB 301 ATCTGGGGCTGGCTATGCTGAGATTGCCAGGCTGACACTCCCTGGAGCCTTCTTT 360  
 QY 178 AspSerLeuValLysGlnThrHisValProAsnLeuPheSerLeuGlnLeuCySerGlyAla 197  
 DB 361 GACTCTCTGTAAAGAGAGAGCCAGTTCCCAACTCTTCTCCCTGACGCTTTGGGTGCT 420  
 QY 198 GlyPheProLeuAsnGlnSerGluValLeuAlaSerValGlyLysSerMetIleIleGly 217  
 DB 421 GGCTTCCCTCAACAGCTGTAAGTGTGGCTCTGTCCGAGGAGCATGATCATTTGA 480  
 QY 218 GlyIleAspHisSerLeuTyrThrGlnLysSerLeuTrpTyrThrProIleArgArgGluTrp 237  
 DB 481 GGATGACCACTGCTGTACAGAGAGCTCTGTGATACCCCATCCGGCGGAGTGG 540

QY 238 TyrTyrGluValIleIleValArgValGluIleAsnGlyGlnAspLeuLysMetAspCys 257  
 DB 541 TATTATGAGGTGATCATTTGCGGGGTGAGATCAATGACAGAGATCTGAAAATGGACTGC 600  
 QY 258 LysGluTyrAsnTyrAspLysSerIleValAspSerGlyThrThrAsnLeuArgLeuPro 277  
 DB 601 AAGAGATCAACTATGACCAAGAGCATTTGTGACAGTGGCACCACTTCGTTGGCC 660  
 QY 278 LysLysVal-PheGluAlaAlaValLysSerIleLeuValAspSerThrGlnLysPhe 297  
 DB 661 AAGAAAGTGTGTAAGCTGAGTCAATTCATCAAGGAGAGCTTCCTCAC-GAGGAGTT 719  
 QY 297 eProAspGlyPheThrLeuGlyGluGlnLeuValCysTrpGlnAlaGlyThrThrProTr 317  
 DB 720 CCTGTATGTTCTGGCTAGAGAGAGCACTGTGTGTCTGGAGAGGAGCACACCCCTTG 779  
 QY 317 PAsnIlePhe-ProValIleSerLeuTyrLeuMetGly---GluValThrAsnGln-Ser 335  
 DB 780 GAACATTTTCCCACTATCTCTCACTTAACCTAAGGGGTAGAGTTACCAACCAAGTCC 839  
 QY 336 Phe-ArgIleThrIleLeuProGln--GlnTyrLeuArg 347  
 DB 840 TTCCCGATCAACATCCCTCCCGAGAGCAATTAACCTGCGG 879

Search completed: August 22, 2005, 05:04:19  
 Job time : 5810 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using SW model

Run on: August 21, 2005, 09:05:38 ; Search time 59 Seconds  
(without alignments)  
706.133 Million cell updates/sec

Title: US-10-726-967a-84

Perfect score: 2288

Sequence: 1 TQHGIRLPRLRSGLGAPLGL.....FTVLDMEDCGYNIPQTEEST 433

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2267	99.1	501	2 A59090	aspartic proteinase
2	311	13.6	384	2 JC7574	pepsinogen A - Afr
3	311	13.6	387	2 B38302	pepsin (EC 3.4.23.
4	308	13.5	385	2 JC7575	pepsinogen A - bul
5	304.5	13.3	383	2 JC7573	pepsinogen C - Afr
6	302.5	13.2	388	1 S19682	pepsin A (EC 3.4.2
7	302	13.2	384	2 A39314	gastric pepsin (EC 3.4
8	301	13.2	382	1 PECH	pepsin A (EC 3.4.2
9	299.5	13.1	396	2 A34401	cathepsin E (EC 3.
10	298.5	13.0	383	2 A41443	pepsin (EC 3.4.23.
11	296	12.9	387	2 C38302	cathepsin B (EC 3.
12	296	12.9	391	2 A43356	cathepsin D (EC 3.
13	296	12.9	407	1 KHRD	cathepsin D (EC 3.
14	295.5	12.9	412	1 KHRD	pepsin (EC 3.4.23.
15	295	12.9	387	2 D38302	pepsin A (EC 3.4.2
16	289.5	12.7	388	1 S19684	cathepsin D (EC 3.
17	287.5	12.6	410	1 KHRD	hypothetical prote
18	287.5	12.6	444	2 T24204	pepsin (EC 3.4.23.
19	286	12.5	387	2 E38302	cathepsin B (EC 3.
20	285.5	12.5	398	2 S66405	chymosin (EC 3.4.2
21	285	12.5	380	2 I51185	chymosin D (EC 3.
22	285	12.5	388	2 I51185	cathepsin D (EC 3.
23	283.5	12.4	388	1 PEMQAR	pepsin A (EC 3.4.2
24	282.5	12.3	389	1 PEMQAR	pepsin C (EC 3.4.2
25	282	12.3	389	2 JEO371	chymosin (EC 3.4.2
26	281.5	12.3	381	1 CMSHB	pepsin A (EC 3.4.2
27	281.5	12.3	388	2 PEMQAR	pepsin A (EC 3.4.2
28	281.5	12.3	388	2 A30142	pepsin A (EC 3.4.2
29	279.5	12.2	388	2 B30142	pepsin A (EC 3.4.2

## ALIGNMENTS

30	278.5	12.2	386	1	PEPG
31	278	12.2	387	2	JC7245
32	277.5	12.0	396	2	S36865
33	273.5	12.0	381	1	CMBO
34	270.5	11.8	377	1	PEMQCJ
35	270.5	11.8	389	2	A38302
36	270	11.8	376	2	I45856
37	268.5	11.7	344	1	KHRPD
38	267.5	11.7	381	2	JC7247
39	267	11.7	380	2	S03433
40	266	11.6	405	2	A25379
41	265.5	11.6	396	2	T47207
42	263.5	11.5	388	2	JC7246
43	262.5	11.5	394	2	B43356
44	261.5	11.4	387	2	A45117
45	261.5	11.4	388	2	A29937

pepsin A (EC 3.4.2  
pepsinogen A - com  
cathepsin E (EC 3.  
chymosin (EC 3.4.2  
gastric pepsin (EC 3.4  
pepsin (EC 3.4.23.  
aspartic proteinase  
cathepsin D (EC 3.  
prochymosin - comm  
candidapepsin (EC  
bactaropepsin (EC  
aspartic proteinase  
pepsinogen C - com  
gastric pepsin (EC 3.4  
aspartic proteinase  
gastric pepsin (EC 3.4

Db 262 VIVAVEINGODLRKMDCKEYVDSIVSGTTNLRPKKVEAAVKSIIKAASTEKPEDG 321  
QY 301 FVLGEOLVCMQAGTTPMNI FVVISLYLMGEVTVNSFRITILPQOYLRPVEDVATSDDCY 360  
Db 322 FVLGEOLVCMQAGTTPMNI FVVISLYLMGEVTVNSFRITILPQOYLRPVEDVATSDDCY 381  
QY 361 KPAISQSGTGMVAVMEGFYVVPDRARRKIGFAVSAACHVHDEFRTAAVEGPFVTLDM 420  
Db 382 KPAISQSGTGMVAVMEGFYVVPDRARRKIGFAVSAACHVHDEFRTAAVEGPFVTLDM 441  
QY 421 DCGYNIPQTDST 433  
Db 442 DCGYNIPQTDST 454

## RESULT 2

pepsinogen A - African clawed frog  
C/Species: Xenopus laevis (African clawed frog)  
C/Date: 30-Jun-2001 #sequence\_revision 30-Jun-2001 #text\_change 09-Jul-2004  
C/Accession: J07574, PC7119  
R/Ikuzawa, M.; Inokuchi, T.; Kobayashi, K.; Yasumasa, S.  
J. Biochem. 129, 147-153, 2001  
A/Title: Amphibian pepsinogens: Purification and characterization of Xenopus pepsinogens  
A/Reference number: J07573; MUID:21064922; PMID:11134969  
A/Contents: Stomach  
A/Accession: J07574  
A/Molecule type: mRNA  
A/Residues: 1-384 <IKU>  
A/Cross-references: UNIPROT:Q9DEC2; DDBJ:AB045380  
A/Accession: PC7119  
A/Molecule type: protein  
A/Residues: 16-35:57-76 <IK2>  
C/Comment: This protein is a zymogen for gastric aspartic proteinase, with pepsin-like  
A/Gene: P9A  
C/Superfamily: pepsin  
C/Keywords: stomach; zymogen

Query Match 13.6%; Score 311; DB 2; Length 384;  
Best Local Similarity 25.1%; Pred. No. 2e-17;  
Matches 107; Conservative 69; Mismatches 154; Indels 96; Gaps 18;  
QY 5 IRLPLRSLGGAFLGLEINLETDEBEPGRGRGFVE-----MVDNLR 47  
Db 17 VKVPLRK-----ESFRNRQRGLGLGVILKKVPNPAKVPPTLAQSAETLQ 65  
QY 48 GKSQGYVEMTVGSPQTLNLTVDGSSNFAVGAAPHFL-----HRYQRQLSSTY 100  
Db 66 NTMDLEYGTISIGPPQEFVIFPDGSANLWV---PSYCSSQACSNHNFNPQSSSTF 122  
QY 101 RDLRKGVVPTQKKEGELGTDLVSIPHPNVTVRANATTESDK-FINGSMWEIL 159  
Db 123 QATNTPVSIQYGTGMSGLGYDTLQV---GNIQISNQMGLSESEPSGFLYSPFDGL 179  
QY 160 GLAYAEIARPDLSLEPFDSLVKQTHVP-NFSLQLCGAGPLNQSEVLASVGSMTIG 218  
Db 180 GLAFPSIA--SSQATPVPFNWMSQGLIPQLNLSVYLSSDG-----QTGSVYLFEG 227  
QY 219 IDHSLYTGSIMWTPIRREWYEVIIIVRVEINGODL--KMDCKEYVDSIYDSGTTNLR 276  
Db 228 VDNSSYSSGLNWNPLTAETIWTITLDSVISINGVYIACQSC--QATVDGTSIMWTG 281  
QY 277 PKRVEAAVKSIIKAASTEKPEDGFVLGEOLV-CMQAGTTPMNI FVVISLYLMGEVTVNS 335  
Db 282 PSTPI-ANIQNYIGASQDSN-----GQYVINCNNISNMFTIYF----- 318  
QY 336 FRITLLPQOYL-RPVEDVATSDDCYK-FAISQSGTGT---VNGAVIMEGFYVVPDRAR 389  
Db 319 ---TINVGQYPLSPASVYRONQCGSSGFGQAMNLPITNSGLMILGADVFIROYFTVFDNRAN 375  
QY 390 KRIGFA 395

Db 376 NYVATA 381

## RESULT 3

pepsin (EC 3.4.23.-) II-1 precursor - rabbit  
C/Species: Oryctolagus cuniculus (domestic rabbit)  
C/Date: 14-Jun-1991 #sequence\_revision 20-Sep-1991 #text\_change 16-Aug-2004  
C/Accession: B38302  
R/Kagiyama, T.; Tanabe, K.; Koiwai, O.  
J. Biol. Chem. 265, 17031-17038, 1990  
A/Title: Structure and development of rabbit pepsinogens. Stage-specific zymogens, nucle  
A/Reference number: A38302; MUID:91009127; PMID:2129536  
A/Accession: B38302  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-387 <KAG>  
A/Cross-references: UNIPROT:P28712; GB:M59235; GB:J05638  
C/Superfamily: Pepsin  
C/Keywords: aspartic proteinase; hydrolase; phosphoprotein; protein digestion

Query Match 13.6%; Score 311; DB 2; Length 387;  
Best Local Similarity 26.6%; Pred. No. 2e-17;  
Matches 105; Conservative 72; Mismatches 145; Indels 72; Gaps 17;

QY 23 NLETDEBEPGRGRGFVEM-VNLRKSGQGYVEMTVGSPQTLNLTVDGSSNFAVG 81  
Db 48 NLATKYPFKE-----TFASVSTSLNLYDAEYFGTISICTPQDEFVTFDITGSSNLMWP 102  
QY 82 ----AAHPLRLRYQRQLSSTYRDLRKGVVPTQKKEGELGTDLVSIPHPNVTVA 137  
Db 103 STYCSLACFLHRRFNDSTFQATSELSITYGSGMTGILGYDTKVV---GNIEDTN 159  
QY 138 NIAITESD---KEFINGSMWEILGLAYAEIARPDLSLEPFDSLVKQTHV-PNIFSLQ 193  
Db 160 QIFSLKTEBGTITFV--APFDGILGLAIPTSIASAT--PVFDNMNGLVSEDLFSVY 215  
QY 194 LCGAGPLNQSEVLASVGSMTIGIDHSLYTGSIMWTPIRREWYEVIIIVRVEINGODL 253  
Db 216 LSSNG-----BKGSVMWFGGIDSSYVTSGLNWNVPSHEGWQIMTDSITNGERTI 265  
QY 254 KM--DCKEYVDSIYDSGTTNLRPKKVEAAVKSIIKAASTEKPEDGFVLGEOLV-CW 310  
Db 266 ACADSC-----QAVDTGISLAGFTSAISKIQTYSIGSKVL-----LGENIISCS 311  
QY 311 QACGTFMNI FVVISLYLMGEVTVNSFRITILPQOYLRPVED--VATSODDC--YKFAISQ 366  
Db 312 AIDSLPDIVF-----TINNVQPLPASAVIILKEDDDCLSGFGQAMLD 353  
QY 367 SSTGT--VNGAVIMEGFYVVPDRARRKIGFAVSA 398  
Db 354 TSYGELMILGADVFIROYFTVFDNRANQVGLAAA 387

## RESULT 4

pepsinogen A - bullfrog  
C/Species: Rana catesbeiana (bullfrog)  
C/Date: 30-Jun-2001 #sequence\_revision 30-Jun-2001 #text\_change 09-Jul-2004  
C/Accession: J07575  
R/Ikuzawa, M.; Inokuchi, T.; Kobayashi, K.; Yasumasa, S.  
J. Biochem. 129, 147-153, 2001  
A/Title: Amphibian pepsinogens: Purification and characterization of Xenopus pepsinogens  
A/Reference number: J07573; MUID:21064922; PMID:11134969  
A/Contents: Stomach  
A/Accession: J07575  
A/Molecule type: mRNA  
A/Residues: 1-385 <IKU>  
A/Cross-references: UNIPROT:Q9DEC4; DDBJ:AB045376  
C/Comment: This protein is a zymogen for gastric aspartic proteinase, with pepsin-like  
A/Gene: P9A  
C/Superfamily: pepsin

C:Keywords: stomach; zymogen

Query Match 13.5%; Score 308; DB 2; Length 385;  
Best Local Similarity 27.8%; Pred. No. 3.5e-17;  
Matches 99; Conservative 64; Mismatches 133; Indels 60; Gaps 15;

QY 54 YVENVTVSSPPTLNLVDTGSSNFAVG---AAPHPLRIHYRQQLSTYDILAKGVV 109  
DB 73 YFGTISIGTPPQSFVIVFDPTGSSNLMWPSVSCSPACTNHMFNPQOSTFQATNTVPASI 132  
QY 110 PYTQKMGEGELTDVLSIPHGPNVTVRANIAITAESDK-FPINGSNMBGILGLAYAEIAR 168  
DB 133 QGTGSMGSGFLGIDTVQV---GNIQITNQIFGLSGSEPSFLYSPFDGILLGLAFPSLA- 188  
QY 169 PDDSLPEFPDLSVKQTHVP-NLPSLQLCGAGFPLNQSEVLAVSGSMIIGIDHSILYGS 227  
DB 169 -SSQATPVPDMMNQGLIQDLPFSVYLSGQ---QS-----GSFVLGGVDTSTYTCN 237  
QY 228 LMVTPIRREWYEVIIIVVEINGODLKM--DCKENYDKSYDSGTTNLRPKVFEAAV 285  
DB 238 LMVPLTAEVYMQITVDSISIGQVYACSGSC-----SAIVDTGTSLAGP---STPI 287  
QY 266 KSIXAASSTKPPDGFMLGEOLVCMQAGTTPMNIPEVVISLYMGEVTVNSFRITILPOQY 345  
DB 268 ANIQYITGANQDSNGQYV---INCNNISMPYVVF-----TINGVQY 326  
QY 346 LRPVED-VATSODDC---YKPAISQSTGT--VMGAVIMEGFYVFPDRAKIGFA 395  
DB 327 PLPAAVAVRQSQSCSSGQANMLPTSSDLMILGDFRERYVVDRAANNVYAMA 382

#### RESULT 5

JC7573  
pepsinogen C - African clawed frog  
N:Alternate names: progastricin  
C:Species: Xenopus laevis (African clawed frog)  
C:Date: 30-Jun-2001 #sequence\_revision 30-Jun-2001 #text\_change 09-Jul-2004  
C:Accession: JC7573; PC7118  
R:Ikuzawa, M.; Inokuchi, T.; Kobayashi, K.; Yasumasa, S.  
J. Biochem. 129, 147-153, 2001  
A:Title: Amphibian pepsinogens: Purification and characterization of Xenopus pepsinogens  
A:Reference number: JC7573; MUID:21064922; PMID:11134969  
A:Contents: Stomach  
A:Accession: JC7573  
A:Molecule type: mRNA  
A:Residues: 1-383 <IKU>  
A:Cross-references: UNIPROT:Q9DEG3; DDBJ:AB045379  
A:Accession: PC7118  
A:Molecule type: protein  
A:Residues: 17-68 <IK2>  
C:Comment: This protein is a zymogen for gastric aspartic proteinase, with pepsin-like  
C:Genetics:  
A:Gene: Pgc  
C:Superfamily: pepsin  
C:Keywords: stomach; zymogen

Query Match 13.3%; Score 304.5; DB 2; Length 383;  
Best Local Similarity 26.0%; Pred. No. 6.8e-17;  
Matches 112; Conservative 64; Mismatches 140; Indels 115; Gaps 19;

QY 2 QHGIRLPLBSGLGAPLGLINLETDEBEPGRKRSFVEMVNLRGKSGQGYVEMTVG 61  
DB 34 ENGIRKAPL-----VDPATKYNQVATAVEPLSNYDMS---YYGHSISIG 74  
QY 62 SPPTQNLILVDTGSSNFAVGA-----AHPFLRIHYRQQLSTYDILAKGVV 113  
DB 75 TTPQNPLVLPDTSNNLMWASTYCQSOACTNHPL---FVPSQSTYSNNQGFSLQYGT 130  
QY 114 GKKEGELGTDVLSIPHGPNVTVRANIAITAESDKFPIGNSN-----WEGILGLAYAEIAR 168  
DB 133 GSIITGLIGYTVI---QNVAIQOBERGLSETP---GTNPFYAGDFDGLIGLAYPSIA- 182  
QY 169 PDDSLPEFPDLSVKQTHVPNLPSLQLCGAGFPLNQSEVLAVSGSMIIGIDHSILYGS 228

DB 183 -VGAGTTVMQGMQO-----NLNQPI--FGFVLSGQS--SQNGGEVAFGVQDVQVYTSQI 233  
QY 229 WYTPIRREWYEVIIIVVEINGOD--LRKDKENYDKSYDSGTTNLRPKVFEAAV 285  
DB 234 YMTPTVSTYMQIGIQGSPINGQATGWCSCOC-----QALVDTGTSILTPAQVFSLSI 287  
QY 286 KSIXAASSTKPPDGFMLGEOLVCMQAGTTPMNI--FPVISLYLMG-----EVTNQ 334  
DB 288 QSIGAQODN-----GQYVSGS-----NIQNLPTISFTISGVSFPLPRAVYVLO 333  
QY 335 S-----FRITILPOQYLRPEVEDVATSQDDCYKPAISQSGTVMGAVIMEGFYVFPDR 387  
DB 334 SSGYTTGIMPTVLPQSGQGL-----WLDGVFRIREYVSVDL 372  
QY 368 ARKRIQFAVSA 398  
DB 373 GNNQVGFATAA 383

#### RESULT 6

S19682  
pepsin A (EC 3.4.23.1) 4 precursor - Japanese macaque  
N:Alternate names: pepsinogen A isozyme 4  
C:Species: Macaca fuscata (Japanese macaque)  
C:Date: 22-Nov-1993 #sequence\_revision 19-Oct-1995 #text\_change 16-Aug-2004  
C:Accession: S19682; S16065  
R:Kageyama, T.; Tanabe, K.; Koizumi, O.  
Eur. J. Biochem. 202, 205-215, 1991  
A:Title: Development-dependent expression of isozymogens of monkey pepsinogens and struc  
A:Reference number: S19681; MUID:92037645; PMID:1935977  
A:Accession: S19682  
A:Molecule type: mRNA  
A:Residues: 1-388 <KAG>  
A:Cross-references: UNIPROT:P27678; EMBL:X59753; NID:938070; PID:CAA42425.1; PID:938071  
A:Note: parts of sequence, including amino ends of pepsinogen and activation intermediat  
C:Comment: This is a minor component of pepsin at all post-partum stages.  
C:Comment: Although two-step activation is observed, activation is predominantly a one-  
C:Superfamily: Pepsin  
C:Keywords: aspartic proteinase; gastric juice; hydrolyase; phosphoprotein; protein diges  
F:1-15/Domain: signal sequence #status predicted <SIG>  
F:16-388/Product: pepsinogen A 4 #status experimental <APT>  
F:16-62/Domain: activation peptide #status experimental <APPT>  
F:63-388/Product: pepsin A 4 #status experimental <ENZ>  
F:38-39/Cleavage site: Leu-Lys (pepsin) #status experimental  
F:62-63/Cleavage site: Leu-Ile (pepsin) #status experimental  
F:94/277/Active site: Asp #status predicted  
F:107-112,268-272,311-344/Diulfide bonds: #status predicted  
F:130/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 13.2%; Score 302.5; DB 1; Length 388;  
Best Local Similarity 27.6%; Pred. No. 1e-16;  
Matches 107; Conservative 64; Mismatches 133; Indels 83; Gaps 17;

QY 27 DEBEPGRKRSFVEMVNLRGKSGQGYVEMTVGSPPTNLILVDTGSSNFAVGAAPR 86  
DB 64 DQOPLR-----NYLVV-----EYFGTIGIGPAPNFTVVDTGSSNLMV---PSV 105  
QY 87 PL-----HRYRQQLSTYDILRKGVVFPYTGKMBEGELTDVLSIPHGPNVTVRANI 139  
DB 106 YCYSLACMDHNLFNQDSTYRATKTVSITGTGSMGILGIDTVQV---GHSIDTNQI 162  
QY 140 AAITESDK-FPINGSNMBGILGLAYAEIARPPDLSLFPDLSVKQTHV-PNLPSLQLCGA 197  
DB 163 FGLSETPGFVFPAPFDGILLGLAYPSIS--SSGATPVPEDNIMNRLVSQDLFSVYLSAD 220  
QY 198 GFLNQSVELAVSGSMIIGIDHSILYGSLSYTPPIRREWYEVIIIVVEINGODL--KM 255  
DB 221 ---DQS-----GSVYVFGIGDSSYTGSLMWNVPSVSGYQISVDSITMNGKITACK 270  
QY 256 DCKENYDKSYDSGTTNLRPKVFEAAVKSIXAASSTKPPDGFMLGEOLV-CMQAGT 314  
DB 271 GC-----QALVDTGTSILTPSPIANIQSDIGASNSD-----GEMVVSQSAIS 316

QY 315 TPWNIFPISLYMGEVTNQSFRITLPEQYLRPEVDVATSDDCYK-----FAISQSS 368  
 Db 317 LPDVF-----TINQVQYPLPPSAVILIQSSQSSCTSGQGMVPTRESG 358  
 QY 369 TGTWGAIVMEGFYVFPDRARKRIGFA 395  
 Db 359 ELWILGDFIRQYFTVFPDRANNQVGLA 385

## RESULT 7

A39314  
 gastricein (EC 3.4.23.3) precursor - bullfrog  
 C/Species: Rana catesbeiana (bullfrog)  
 C/Date: 19-Jun-1992 #sequence\_revision 19-Jun-1992 #text\_change 16-Aug-2004  
 C/Accession: A39314  
 Rikabe, E.; Tanji, M.; Ichinose, M.; Goto, S.; Miki, K.; Kurokawa, K.; Ito, H.; Kageya  
 J. Biol. Chem. 266, 22436-22443, 1991  
 A/Title: Purification, characterization, and amino acid sequences of pepsinogens and pep  
 A/Reference number: A39314; MUID:92042186; PMID:1939266  
 A/Accession: A39314  
 A/Status: preliminary  
 A/Molecule type: mRNA  
 A/Residues: 1-384 <YAK>  
 A/Cross-references: UNIPROT:Q91322; GB:M73750; NID:g213687; PIDN:AAA49530.1; PID:g213688  
 C/Superfamily: Pepsin  
 C/Keywords: aspartic proteinase; hydrolase; protein digestion

Query Match 13.2%; Score 302; DB 2; Length 384;  
 Best Local Similarity 24.5%; Pred. No. 1.1e-16;  
 Matches 105; Conservative 65; Mismatches 147; Indels 112; Gaps 17;  
 QY 3 HGIRLPRLSGLAGAPLGLINLETDEPERGR--GSFVEMVDNRKSGCGCYVEMTV 60  
 Db 35 HGIRKAV-----VDATKYNNRATRAPEPLANTYMDMSYGEISL 73  
 QY 61 GSPPTQNLINIVDTGSSNFAVGAAPHPL-----HRYVQRQLSSTYRDLRKGVVPTYQ 113  
 Db 74 GTPQNFVLFDTGSSNLWV---PSTYCSQAQCTNHPQNPSSSSYSNQQQFSLQYCT 130  
 QY 114 GKMEGELGDLVSIPIHGPNVTYRANIA-----AITSDFKPFINGSMWEGILGLAYAE 165  
 Db 131 GSITGLIGDTVQI-----QNTAISQGEFGSLVTEPGTNFVY-AQFDGILGLAYPS 180  
 QY 166 IARPDLSLEPPFSDLVKQTHVN-LFSLQLCGAGFPLNQSSEVLASVSGMIIGIDHSLY 224  
 Db 181 IAA--EGATVWQGMQQLINQLPFAFLSQQQNSQ-----GSEVAFGVDQNY 230  
 QY 225 TGSIMTPIRREMYEVIIIVVEINGQD--LMDCKEYNDKSIYDSGTTNLRLPKVF 281  
 Db 231 SQQIWTPTSTBYQIQIGSPSVNGQATGWCSCG-----QGIYDTGTSILTPAQSVF 284  
 QY 282 EAAVKSIRKAASTEKFPDGFMLGEOLV-CWQAGTTPMNIFFVY-----SLYMGVTV 332  
 Db 285 SLMQSIGAQDON-----GQYAVSCSNISQSLPTISFTISGVSPPLPSAAYVLQNS 336  
 QY 333 NO--SFRTITLPQYLRPEVDVATSDDCYKFAISQSSSTGTWGAIVMEGFYVFPDRAR 389  
 Db 337 GYCTIGIMPTVLPSONQGPL-----WILGDFVLRQYISVYDIGN 375  
 QY 390 KRIGFVA 398  
 Db 376 NOVGFRAAA 384

## RESULT 8

pepsin A (EC 3.4.23.1) precursor - chicken  
 N/Alternate names: pepsinogen A  
 C/Species: Gallus gallus (chicken)  
 C/Date: 18-Apr-1984 #sequence\_revision 01-Dec-2000 #text\_change 16-Aug-2004  
 C/Accession: J03370; A00984  
 R/Sakamoto, N.; Saito, H.; Yasugi, S.

Biochem. Biophys. Res. Commun. 250, 420-424, 1998  
 A/Title: Analysis of temporal expression pattern and cis-regulatory sequences of chicken  
 A/Reference number: J03370; MUID:98440813; PMID:9753645  
 A/Accession: J03370

A/Status: preliminary  
 A/Molecule type: mRNA  
 A/Residues: 1-382 <SAK>  
 A/Cross-references: UNIPROT:Q9PRG9; GB:AB025281; NID:g4589837; PIDN:BA476891.1; PID:g4589  
 R/Baudys, M.; Kostka, V.  
 Eur. J. Biochem. 136, 89-99, 1983  
 A/Title: Covalent structure of chicken pepsinogen.  
 A/Reference number: A00984; MUID:84004412; PMID:6617663  
 A/Accession: A00984  
 A/Molecule type: protein  
 A/Residues: 16-87, 'S', '89-382 <BAU>  
 C/Superfamily: Pepsin  
 C/Keywords: aspartic proteinase; gastric juice; glycoprotein; hydrolase; protein digestive  
 F/16-57/Domain: activation peptide #status experimental <APT>  
 F/58-382/Product: pepsin A #status predicted <MAT>  
 F/92,275/Active site: Asp #status predicted  
 F/105-110,266-270,305-338/Disulfide bonds: #status experimental  
 F/128/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 13.2%; Score 301; DB 1; Length 382;  
 Best Local Similarity 24.0%; Pred. No. 1.3e-16;  
 Matches 88; Conservative 68; Mismatches 126; Indels 84; Gaps 13;  
 QY 54 YVEMTVGSPPTQNLINIVDTGSSNFAVGAAPHPL-----HRYVQRQLSSTYRDLRK 106  
 Db 74 YGIGTISGTQPDPTVIFDTGSSNLWV---PSTYCSQAQCTNHPQNPSSSSYSNQQQFSLQYCT 130  
 QY 107 VYVPTQKMEGELGDLVSIPIHGPNVTYRANIAITSDK--FPINGSMWEGILGLAYAE 165  
 Db 131 VYIAYGTGSMGILGYDTAV---SSIDVQNIQFGESETPGSGFFYCNFPGDILGLAFPS 187  
 QY 166 IARPDLSLEPPFSDLVKQTHV-PNLFSLOLCGAGFPLNQSSEVLASVSGMIIGIDHSLY 224  
 Db 188 IS--SSATVTFVDMQMSQHLVAQDLFSVYSKSG-----ETGSEFVLFEGIDPNYT 235  
 QY 225 TGSIMTPIRREMYEVIIIVVEINGQDLK--MDCKEYNDKSIYDSGTTNLRLPKVF 282  
 Db 236 TKGIYVWPLAEFTYMQITMDRVTYGNKVKVACFPTC-----QAIYDTGTSILVMPGCAV 289  
 QY 283 AAVKSIKAASTP-----KFPDGFMLGEOLVWCQAGTTPMNIFFVISLYMGEVTNOS 335  
 Db 290 RIHKDLGVSSDGIISCDISKLDP-----VTFHNGHA----- 322  
 QY 336 FRITITLPQYLRPEVDVATSDDCYKFAISQSSSTGT-----VMGAVIMEGFYVFPDRAR 389  
 Db 323 -----FTLPASAVVINEDESCMLGFENMCTPTELGQWILGDFIRREYVIFDRAN 373  
 QY 390 KRIGFVA 395  
 Db 374 NKVGLS 379

## RESULT 9

A34401  
 cathepsin E (EC 3.4.23.34) precursor - human  
 C/Species: Homo sapiens (man)  
 C/Date: 22-Jun-1990 #sequence\_revision 22-Jun-1990 #text\_change 16-Aug-2004  
 C/Accession: A42038; A34401; S35663; S34467; A34643; B31643  
 R/Azuma, T.; Liu, W.; Vander laan, D.J.; Bowcock, A.M.; Taggart, R.T.  
 J. Biol. Chem. 267, 1609-1614, 1992  
 A/Title: Human gastric cathepsin E gene. Multiple transcripts result from alternative po  
 A/Reference number: A42038; MUID:92112877; PMID:1370478  
 A/Accession: A42038  
 A/Molecule type: DNA  
 A/Residues: 1-396 <AZU>  
 A/Cross-references: UNIPROT:P14091; GB:M84424; GB:M82847; NID:g181203; PIDN:AAA52300.1; I  
 R/Azuma, T.; Fals, G.; Monandas, T.K.; Couvreur, J.M.; Taggart, R.T.  
 J. Biol. Chem. 264, 16748-16753, 1989





```
Matches 97; Conservative 63; Mismatches 135; Indels 66; Gaps 13;
QY 54 YVEKTVGSPPTQTLILVDTGSSNFAVGAAPHF-----LHRYGQRLSSYTRDLRKG 106
Db 75 YFGTISIGPPQDFVITIEDTGSSNLWV---PSTYCSLALCHKRFNEDSSTYQGSRT 131
QY 107 VVVPYTGKMEGLGTDLVSI PHGPNVTVRANIAITSDKFFINGSMWEGILGLAVAEI 166
Db 132 LSTYTGSMGILGIDYTVKVGSLBDTQIFGLSTEPSLFLF--APFDGILLALAPSI 189
QY 167 ARPDLSLEPFPSLVKQTHV--PNLFSIQLCGAGFPLNQSSEVLASVGSMITGGIDHSLYT 225
Db 190 SSSDAT--PVEDNMNMEGLVSGDLFSVYLLSSD-----EKSGLVWFGGIDSSYIT 237
QY 226 GSIWTPIRRMYYEVIIVREINGDLKM--DKEYNYSIYDSTGTTMLRPPKVFEEA 283
Db 238 GSIWNPVSYEGVQIOTDVSINGETIACADSC---QAIYDTGSLTGP---TS 287
QY 284 AVKSIKAASTKEPDPGFWEGLV--CWAQGTTPMNFVYISLVLMGEVTVNSFRITILP 342
Db 288 AISNIQSYIGASK---NLGSENVISCSAIDSLPDIYF-----TING 325
QY 343 QQYLRPEVDVATSDDCYKFAISQSSSTGT-----VNGAVIMEGFVVPDRARRKIGFAV 396
Db 326 IQYPLPASAVYLKEDDCTSGLEGANVDYTGELMILGIDVFIRQYFTVPDRANNQGLAA 385
QY 397 S 397
Db 386 A 386

RESULT 12
A43356
cathepsin E (EC 3.4.23.34) precursor - guinea pig
N:Alternate names: erythrocyte membrane aspartic proteinase; slow-moving proteinase
C:Species: Cavia porcellus (guinea pig)
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 16-Aug-2004
C:Accession: A43356
R:Kageyama, T.; Ichinose, M.; Tanukada, S.; Miki, K.; Kurokawa, K.; Koizumi, O.; Tanji, M.
J. Biol. Chem. 267, 16450-16459, 1992
A:Title: Gastric procathepsin E and progastricin from guinea pig. Purification, molecu
A:Reference number: A43356, MUID:92355614; PMID:1644829
A:Accession: A43356
A:Molecule type: mRNA
A:Residues: 1-391 <NAG>
A:Cross-references: UNIPROT:P25796; GB:W88653; NID:9191294; PIDN:AAA7052.1; PTD:9191295
A:Note: sequence extracted from NCBI backbone (NCBI:110763, NCBI:P.110763)
C:Superfamily: Pepsin
C:Keywords: aspartic proteinase; hydrolase; membrane protein

Query Match 12.9%; Score 296; DB 2; Length 391;
Best Local Similarity 26.9%; Pred. No. 3.4e-16;
Matches 98; Conservative 64; Mismatches 130; Indels 72; Gaps 16;
QY 54 YVEMTVGSPPTQTLILVDTGSSNFAVGA---AHPFLHRYGQRLSSYTRDLRKYVY 109
Db 74 YFGTISIGPPQDFVITIEDTGSSNLWVSVCTSPACQTHPFRHSLSTREVNRSFRI 133
QY 110 PYTGKMEGLGTDLVSI PHGPNVTVRANIAITSDKFFINGSMWEGILGLAVAEIAR 169
Db 134 QYGTSLTIGIIGADQVSV--EGLTVVGGQFGESVQDPRGKTFPH--AFPDGILLALAPSI 190
QY 170 DQSLRPFPSLVKQTHVRLPSQLCGAGFPLNQSSEVLASVGSMITGGIDHSLYTSLW 229
Db 191 -GGVTPVEDNMNAQ---NLVALPM---FSVYMSNPGSGSELTFEGYDPSHFSGSLN 241
QY 230 YTPIRREMYEVIIVREINGDLKMDCKEYNYSIYDSTGTTMLRPPKVFEEAAYSIK 289
Db 242 WVPVTKQAVYQIADLGIQV--DSVMFSE--GCQALYDTGSLTGP---PEKIQIQ 293
*QY 290 AASTTEKPPDGFWEGLV--CWAQGTTPMNFVYISLVLMGEVTVNSFRITILPQQ 344
Db 294 BALGATVYDGEY---SVQC-----ANLNMMLDVT---FLINGVPTILNPTA 333
```

```
QY 345 YLRPEVDVATSDDCYKFAISQSSSTG-----TWGAVIMEGFVVPDRARRK 391
Db 334 Y--TLIDFVDMQV-----STFEGLEIQPPAGFLWLGIDVFIRQYFAVEDRGNR 383
QY 392 IGFA 395
Db 384 VGLA 387

RESULT 13
KRTD
cathepsin D (EC 3.4.23.5) precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 16-Aug-2004
C:Accession: S13111; C1918; Q01177; P00222
R:Birch, N.P.; Loh, Y.P.
Nucleic Acids Res. 18, 6445-6446, 1990
A:Title: Cloning, sequence and expression of rat cathepsin D.
A:Reference number: S13111; MUID:91057150; PMID:2243802
A:Accession: S13111
A:Molecule type: mRNA
A:Residues: 1-407 <BIR>
A:Cross-references: UNIPROT:P24268; EMBL:X54467; NID:955881; PIDN:CAA8349.1; PTD:955882
R:Yonezawa, S.; Takahashi, T.; Wang, X.; Wong, R.N.S.; Hartwick, J.A.; Tang, J.
J. Biol. Chem. 263, 16504-16511, 1988
A:Title: Structures at the proteolytic processing region of cathepsin D.
A:Reference number: A92681; MUID:89034127; PMID:3182800
A:Accession: C31918
A:Molecule type: Protein
A:Residues: 134-152, 'T', 164-170 <YON>
R:Fujita, H.; Tanaka, Y.; Noguchi, Y.; Kono, A.; Himeno, M.; Kato, K.
Biochem. Biophys. Res. Commun. 179, 190-196, 1991
A:Title: Isolation and sequencing of a cDNA clone encoding rat liver lysosomal cathepsin
A:Reference number: Q01177; MUID:91354249; PMID:1883350
A:Accession: Q01177
A:Molecule type: mRNA
A:Residues: 1-14, 'A', 16-204, 'N', 206-261, 'N', 263-407 <FUJ>
A:Accession: P00222
A:Molecule type: Protein
A:Residues: 65-74, 118-127, 165-174 <FUJ>
A:Experimental source: liver
C:Function: Cathepsin D in rat liver lysosome occurs as a mixture of both a single chain f
A:Description: limited specificity endopeptidase
A:Pathway: Intracellular protein degradation
C:Superfamily: Pepsin
C:Keywords: aspartic proteinase; glycoprotein; hydrolase; lysosome; protein degradation
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-64/Domain: propeptide #status predicted <PRC>
F:65-407/Product: cathepsin D 43K single-chain form #status predicted <MAT>
F:65-147/Product: (or 65-165) cathepsin D 12K light chain #status predicted <MA2>
F:118-407/Product: cathepsin D 9K light chain #status predicted <MA4>
F:165-407/Product: (or 166-407) cathepsin D 34K heavy chain #status predicted <MA5>
F:91-160, 110-117, 281-285, 324-361/Disulfide bonds: #status predicted <MA3>
F:97, 290/active site: Asp #status predicted
F:134, 258/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 12.9%; Score 296; DB 1; Length 407;
Best Local Similarity 25.6%; Pred. No. 3.6e-16;
Matches 112; Conservative 77; Mismatches 152; Indels 96; Gaps 19;
QY 5 IRLPL-----SGGAGPLG-----EINLEDEPPEPRGRGSFVENVDILRG 48
Db 22 IRLPLKRTSIRRTWTEVGSVEDILKGPITKYSWMSPTKEP-----VSEILKN 73
QY 49 KSGGQYVEMTVGSPPTQTLILVDTGSSNFAVGAAPHF-----HRYGQRLSSYTRD 102
Db 74 YLDAQYGEIGIPIPGQCFIVFPDGSNLWVPSIHCKLDIACWVHHKNSDKSSTYVK 133
QY 103 LKRGVYVYTGKMEGLGTDLVSI PHGPNVTVRANIAITSDKFF-----ING 152
```

Db 134 NCTSPDIHSGSLGSLYSDQTVSVF-----CKSDLGKIKVEKQIFGATKQPGVFIA 187

Qy 153 SNMEGILGLAVAEIARPDLSLEPPDLSLVKQTHV-PNLFSLQICGAGFPLNQSEVLASVG 211

Db 188 ARPDGLMGKGYFIS--VAKVLPVPDNLKQKQKLVENKITS-----FVLNR-DPTGQGG 237

Qy 212 GSMITGIDHSLYTGSLMTTPRREMYEVIIVRVAINODLKMCKENYDKSIYDST 271

Db 238 GELMLGDTSRYYKHGELSYLVNTRKAYQVHMDQLEV-GSELTLC-CK--GGCEAIYDTGT 293

Qy 272 TULRLPKKYFBAVVSIRKASSTKEKPPDGMWLGEOLV-CMOAGTTPMNIIPVISTLYMGE 330

Db 294 SLVGVDEVKELQKATIGAVPLIQ-----GEYVTPCEKVS-----LPITFLKLGQ 340

Qy 331 VTNQSFRTITLPQOYLRPVEDVATSCDDCYKFAIS-----QSTGTWAGAVIMEGF 381

Db 341 -----NVELHPKRTILKVSQAGKT-----ICLSGFMGMDIPPPSGPLMILGDVITGCV 388

Qy 382 YVFDPRARKRIGFAVSA 398

Db 389 YTVFDRBYNRVGFAPKAA 405

RESULT 14

KRHND

cathepsin D (EC 3.4.23.5) precursor [validated] - human

N/Alternate names: preprocathepsin D

C/Species: Homo sapiens (man)

C/Date: 28-Dec-1987 #sequence revision 28-Dec-1987 #text\_change 16-Aug-2004

C/Accession: A25771; S30749; PC2066; I59236; I57716

R/Paut, P.L.; Kornfeld, S.; Chirgwin, J.M.

Proc. Natl. Acad. Sci. U.S.A. 82, 4910-4914, 1985

A/Title: Cloning and sequence analysis of cDNA for human cathepsin D.

A/Reference number: A25771; MUID:85270436; PMID:3927292

A/Accession: A25771

A/Molecule type: mRNA

A/Residues: 1-412 <PAU>

A/Cross-references: UNIPROT:P07339; EMBL:M11233; NID:G181179; PID:MA59529.1; PID:G1811

Nucleic Acids Res. 15, 3773-3786, 1987

A/Title: Oestrogen regulates cathepsin D mRNA levels in oestrogen responsive human breast

A/Reference number: S30749; MUID:87231068; PMID:3588310

A/Accession: S30749

A/Molecule type: mRNA

A/Residues: 1-412 <WES>

A/Cross-references: EMBL:X05344; NID:G29677; PID:CAA28955.1; PID:G29678

R/May, F.E.B.; Smith, D.J.; Westley, B.R.

Gene 134, 277-282, 1993

A/Title: The human cathepsin D-encoding gene is transcribed from an estrogen-regulated a

A/Reference number: PC2066; MUID:94085791; PMID:8262386

A/Accession: PC2066

A/Molecule type: DNA

A/Residues: 1-23 <MAY>

A/Cross-references: GB:LI2980; NID:G291930; PIDN:AAA16314.1; PID:G455429

A/Experimental source: MCF-7 cell

R/Cavallier, V.; Augereau, P.; Rochefort, H.

Proc. Natl. Acad. Sci. U.S.A. 90, 203-207, 1993

A/Title: Cathepsin D gene is controlled by a mixed promoter, and estrogens stimulate onl

A/Reference number: I59236; MUID:93126342; PMID:8419924

A/Accession: I59236

A/Status: translation not shown; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-22 <CAV1>

A/Cross-references: GB:S52557; NID:G263124; PIDN:AA13868.1; PID:G4261568

R/Augerreau, P.; Miralles, P.; Cavallier, V.; Gaudelot, C.; Parker, M.; Rochefort, H.

Mol. Endocrinol. 8, 693-703, 1994

A/Title: Characterization of the proximal estrogen-responsive element of human cathepsin

A/Reference number: I57716; MUID:95021301; PMID:7935485

A/Accession: I57716

A/Status: translation not shown; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-22 <CAV2>

A/Cross-references: GB:S74689; NID:G786350; PIDN:AA14156.1; PID:G4261856

R/Baldwin, E.T.; Bhat, T.N.; Guinik, S.; Erickson, J.W.

submitted to the Brookhaven Protein Data Bank, April 1993

A/Reference number: A51839; PDB:1LVA

A/Contents: annotation; X-ray crystallography, 2.5 angstroms, residues 65-161,170-241

R/Baldwin, E.T.; Bhat, T.N.; Guinik, S.; Erickson, J.W.

submitted to the Brookhaven Protein Data Bank, April 1993

A/Reference number: A51840; PDB:1LVB

A/Contents: annotation; X-ray crystallography, 2.5 angstroms, with inhibitor residues 65

R/Baldwin, E.T.; Bhat, T.N.; Guinik, S.; Hosur, M.V.; Sowder II, R.C.; Cachau, R.E.; Col

Proc. Natl. Acad. Sci. U.S.A. 90, 6796-6800, 1993

A/Title: Crystal structures of native and inhibited forms of human cathepsin D: implicat

A/Reference number: A48229; MUID:93342076; PMID:8393577

A/Contents: annotation; X-ray crystallography, 2.5 angstroms

C/Comment: Cathepsin D is a ubiquitous lysosomal proteinase.

C/Comment: In addition to the propeptide, residues 163-168 and 411-412 are proteolytical

C/Comment: The carbohydrate bound to 134-Asn contains a mannose-6-phosphate that is bou

C/Genetics:

A/Gene: GDB:CTSD

A/Cross-references: GDB:120512; OMIM:116840

A/Map position: 11p15.5-11p15.5

C/Function:

A/Description: limited specificity endopeptidase

A/Pathway: intracellular protein degradation

C/Superfamily: Pepsin

C/Keywords: aspartic proteinase; glycoprotein; hydrolase; lysosome; protein degradation

F/1-20/Domain: signal sequence #status predicted <SIG>

F/21-64/Domain: propeptide #status predicted <PRO>

F/65-162,169-410/Product: cathepsin D #status experimental <MAT>

F/167-329-356/Region: phosphotransferase recognition

F/91-160,110-117,286-290,329-366/Diulfide bonds: #status experimental

F/97,295/Active site: Asp #status experimental

F/134,263/Binding site: carbohydrate (asn) (covalent) #status experimental

Query Match 12.9%; Score 295.5; DB 1; Length 412;

Best Local Similarity 28.5%; Pred. No. 4e-16;

Matches 105; Conservative 58; Mismatches 146; Indels 59; Gaps 15;

Qy 54 YVEKTVSGSPPTLNIIVDTGSSNFAVGAARHPF-----HRYVQRQLSTYRDLRKGV 107

Db 79 YVGEIGIGIPPCQCFYVVDPTGSSNFWPSHCKLDIACWHHKNYSDKSTYVKNQTSF 138

Qy 108 YVPYQGMKEGELGTDLVSIIP-----HGPNTVNRANIAITESDKFFINGSNMGEI 158

Db 139 DIHSGSLGSLGSDQTVSVPCQASASALGCVKVERQVPEARKQPOITIAKFDPI 198

Qy 159 LGLAVAEIARPDLSLEPPDLSLVKQTHV-PNLFSLQICGAGFPLNQSEVLASVGSMITG 217

Db 199 LGMAYPRIS--VNNVLPVPDNLMOQKLVQDNIFSYL-----SRDDAQPGGELMIG 248

Qy 218 GIDHSLYTGSLMYTPPIREMYEVIIVRVEI-NGODLKMCKENYDKSIYDSTTNLRL 276

Db 249 GTDSKYYKGLSLYLVNTRKAYQVHMDQVEVASGLTL---CKE--GCEAIYDTGSLMWG 303

Qy 277 PKKVFPAVKSIKKAASSTKEKPPDGMWLGEOLV-CMOAGTTPMNIIPVISTLYMGEVTNOS 335

Db 304 PDEVERELKAIKAGAPLID-----GEYVTPCEKVS-----LPATFLKLG-----KG 346

Qy 336 FTITLPQOYLRPVEDVATSCDDCYKFAISQ-----SSTGTWAGAVIMEGFVYVDPARK 390

Db 347 YKLS--PEYITLKVSGAGTTL--CLSGFMGMDIPPPSGPLMILGDVIFGRYTVFDRDN 402

Qy 391 RIGFAVSA 398

Db 403 RVGFARAA 410

RESULT 15

D38302

pepsin (EC 3.4.23.-) II-4 precursor - rabbit

C/Species: Oryctolagus cuniculus (domestic rabbit)

C/Date: 14-Jun-1991 #sequence\_revision 20-Sep-1991 #text\_change 16-Aug-2004

C/Accession: D38302

R/Kageyama, T.; Tanabe, K.; Koiwai, O.

J. Biol. Chem. 265, 17031-17038, 1990

**A/Title:** Structure and development of rabbit pepsinogens. Stage-specific zymogens, nucle

A;Reference number: A38302; MUID:91009127; PMID:2129536

A;Accession: D38302

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-387 <KAG>

A:Cross-references: UNIPROT: P28713: GB:M59235: GB:T056318  
M/REBASE: 1-38 / <NAGS

C: Superfamily: Penat

**C;Keywords:** aspartic proteinase; hydrolase; phosphoprotein; protein digestion

Query Match	12.9%;	Score 295;	DB 2;	Length 387;
-------------	--------	------------	-------	-------------

Best Local Similarity	Pred. No.
26.18	4.1e-16

Matches 97; Conservative 65; Mismatches 123; Indels 86; Gaps 14;

```

QY      54  YYYWNTVGSPPOTLNLIVDTGSSNFVGAAPHP-----LHRKYQLSTTYRLRK 106
Db      75  YFGTISIGTPPOQFOFTVIFPTGSSNLVW---PSTYCSLALHKRFNEDBSTYQTSHT 131
QY      107  VVVPYTGKKEGELGTDLVLSIPHGPNVTVRANIAITESDKPF-----INGSNME 156
Db      132  LSTTYGGSMTGLGTDV-----KVGSEDTNQIFGLSKTEBPGLTLPAPFD 179
QY      157  GILGLAVAIARDDSLBPPFDSLVKQTHV-ENLPSLQICGAGPPLNQSEVLASVGSMT 215
Db      180  GILGLAVPSISSDAT-PVFDNMNMEGLVSGDLPFSVYSSDD-----EKGSVM 227
QY      216  IGGIDSLVYGLSLWYFPPIREMYEVYIIVRIVINQDLKM--DCKEYVYDKSYDSTTN 273
Db      228  FGGIDSSYYTGSINWVPVSEGVQIYMDVSINETIACADSC-----QALVDGTSL 281
QY      274  LRLPKVFEAAVGSIAAASSTEFKPDGFWLGEOLV-CWQAGTTPMNIFFVSLYLMGEVT 332
Db      282  LTGP-----TSAINISISYGASK---NLGSGNVISCAISLSDIYF----- 321
QY      333  NQSFRTIILPQOYLREVEDVATSODDCYKFAISQSGTG-----VNGAVIMEGFVVPD 386
Db      322  -----TINGIQYPLPASAYILKEDDCTSGLEGNNVDYTLGELMLIGDVFIRQYFVVP 375
QY      387  RARRKIGFAVS 397
Db      376  RANNQGLAA 386

```

Search completed: August 22, 2005, 00:25:10

Job time : 61 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

## OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 22, 2005, 00:54:10 ; Search time 1103 Seconds

(without alignments)  
2550.468 Million cell updates/sec

Title: US-10-726-967a-84

Perfect score: 2288  
Sequence: 1 TQHGIRLPRLSGLGAPLGL.....FTYLMEDCGINIPQTDST 433Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 7316285 seqs, 3248459403 residues

Total number of hits satisfying chosen parameters: 14632570

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xlp  
-O=/cg2n2.1/USPTO.spool.p/US1072667.r/unat.19082005.201241.29685/app.query.fasta\_1.583  
-DB=Published Applications NA -QWMT=faaap -SUFFIX=crpb -MIMATCH=0.1  
-LOOPL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsu62  
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=pcr -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pcr -NORM=ext -HEADSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USPR=US1072667 @CGN 1.1 723 @runat.19082005.201241.29685  
-NCPU=6 -ICPU=3 -NO MAP -LARGEIOBRY -NEG SCORES=0 -NAT -DSPBIOCK=100  
-LONGLOG -DEV TIMEOUT=120 -MAIN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FPAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:\*

1: /cg2n2\_6/prodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cg2n2\_6/prodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cg2n2\_6/prodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cg2n2\_6/prodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
5: /cg2n2\_6/prodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*  
6: /cg2n2\_6/prodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
7: /cg2n2\_6/prodata/1/pubpna/US08\_PUBCOMB.seq:\*  
8: /cg2n2\_6/prodata/1/pubpna/US09\_PUBCOMB.seq:\*  
9: /cg2n2\_6/prodata/1/pubpna/US09B\_PUBCOMB.seq:\*  
10: /cg2n2\_6/prodata/1/pubpna/US09C\_PUBCOMB.seq:\*  
11: /cg2n2\_6/prodata/1/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cg2n2\_6/prodata/1/pubpna/US09C\_PUBCOMB.seq:\*  
13: /cg2n2\_6/prodata/1/pubpna/US10\_PUBCOMB.seq:\*  
14: /cg2n2\_6/prodata/1/pubpna/US10C\_PUBCOMB.seq:\*  
15: /cg2n2\_6/prodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
16: /cg2n2\_6/prodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
17: /cg2n2\_6/prodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
18: /cg2n2\_6/prodata/1/pubpna/US10C\_PUBCOMB.seq:\*  
19: /cg2n2\_6/prodata/1/pubpna/US10C\_PUBCOMB.seq:\*  
20: /cg2n2\_6/prodata/1/pubpna/US10C\_PUBCOMB.seq:\*  
21: /cg2n2\_6/prodata/1/pubpna/US10C\_PUBCOMB.seq:\*  
22: /cg2n2\_6/prodata/1/pubpna/US10C\_PUBCOMB.seq:\*  
23: /cg2n2\_6/prodata/1/pubpna/US11\_PUBCOMB.seq:\*  
24: /cg2n2\_6/prodata/1/pubpna/US11\_PUBCOMB.seq:\*  
25: /cg2n2\_6/prodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
26: /cg2n2\_6/prodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2267	99.1	1371	19 US-10-281-092-7	Sequence 7, Appli
2	2267	99.1	1503	24 US-11-089-918-1	Sequence 1, Appli
3	2267	99.1	1503	24 US-11-090-866-1	Sequence 1, Appli
4	2267	99.1	1503	24 US-11-069-377-1	Sequence 1, Appli
5	2267	99.1	1503	24 US-11-090-872-1	Sequence 1, Appli
6	2267	99.1	1506	15 US-10-214-932-103	Sequence 5, Appli
7	2267	99.1	1506	19 US-10-281-092-5	Sequence 6, Appli
8	2267	99.1	1527	19 US-10-275-339A-6	Sequence 1, Appli
9	2267	99.1	1542	15 US-10-322-684-1	Sequence 1, Appli
10	2267	99.1	2070	18 US-10-652-927-3	Sequence 3, Appli
11	2267	99.1	2070	18 US-10-652-830-3	Sequence 3, Appli
12	2267	99.1	2070	19 US-10-652-045-3	Sequence 3, Appli
13	2267	99.1	2070	21 US-10-940-867-3	Sequence 3, Appli
14	2267	99.1	2348	24 US-11-089-918-4	Sequence 44, Appli
15	2267	99.1	2348	24 US-11-090-866-4	Sequence 44, Appli
16	2267	99.1	2348	24 US-11-069-377-4	Sequence 44, Appli
17	2267	99.1	2348	24 US-11-090-872-4	Sequence 44, Appli
18	2267	99.1	2526	19 US-10-466-358-1	Sequence 1, Appli
19	2267	99.1	2526	22 US-10-466-391A-1	Sequence 1, Appli
20	2267	99.1	3252	9 US-09-796-264-1	Sequence 1, Appli
21	2267	99.1	3252	9 US-09-845-226-1	Sequence 1, Appli
22	2267	99.1	3252	9 US-09-795-903A-1	Sequence 1, Appli
23	2267	99.1	3252	14 US-10-032-818-1	Sequence 1, Appli
24	2267	99.1	3252	19 US-10-820-953-1	Sequence 1, Appli
25	2267	99.1	3252	20 US-10-773-754-1	Sequence 1, Appli
26	2267	99.1	5832	19 US-10-721-693-18	Sequence 18, Appli
27	2267	99.1	5832	20 US-10-852-997-18	Sequence 18, Appli
28	2267	99.1	5876	20 US-10-723-860-5006	Sequence 5006, Ap
29	2267	99.1	5878	17 US-10-159-942-4	Sequence 4, Appli
30	2267	99.1	5878	20 US-10-723-860-384	Sequence 284, App
31	2267	99.1	5878	21 US-10-956-157-1778	Sequence 1778, Ap
32	2267	99.1	16080	24 US-11-089-918-48	Sequence 48, Appli
33	2267	99.1	16080	24 US-11-090-866-48	Sequence 48, Appli
34	2267	99.1	16080	24 US-11-069-377-48	Sequence 48, Appli
35	2267	99.1	16080	24 US-11-090-872-48	Sequence 48, Appli
36	2262	98.9	1368	18 US-10-627-473-1	Sequence 1, Appli
37	2262	98.9	2070	9 US-09-794-927-3	Sequence 3, Appli
38	2262	98.9	2070	9 US-09-795-847-3	Sequence 3, Appli
39	2262	98.9	2070	9 US-09-794-743-3	Sequence 3, Appli
40	2262	98.9	2070	9 US-09-794-743-3	Sequence 3, Appli
41	2262	98.9	2070	9 US-09-794-925-3	Sequence 3, Appli
42	2262	98.9	2070	9 US-09-681-443-3	Sequence 3, Appli
43	2262	98.9	2070	10 US-09-908-943A-1	Sequence 1, Appli
44	2262	98.9	2070	10 US-09-869-414-3	Sequence 3, Appli
45	2262	98.9	2070	10 US-09-548-366-3	Sequence 3, Appli

## ALIGNMENTS

RESULT 1  
US-10-281-092-7  
; Sequence 7, Application US/10281092  
; Publication No. US20040121947A1  
; GENERAL INFORMATION:  
; APPLICANT: Ghosh, Arun K.  
; APPLICANT: Tang, Jordan J.N.  
; APPLICANT: Bilcer, Geoffrey  
; APPLICANT: Chang, Manpin  
; APPLICANT: Hong, Lin  
; APPLICANT: Koelsch, Gerald E.  
; APPLICANT: Loy, Jeffrey A.  
; APPLICANT: Turner, Robert T., III  
; APPLICANT: Devanumadrum, Thippeswamy  
; TITLE OF INVENTION: COMPOUNDS WHICH INHIBIT BETA-SECRETASE  
; ACTIVITY AND METHODS OF USE THEREOF

```

; FILE REFERENCE: 2932.1001-004
; CURRENT APPLICATION NUMBER: US/10/281.092
; PRIOR FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: US 10/032, 818
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: PCT US01/50826
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: US 60/258, 705
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: US 60/275, 756
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: US 60/335, 952
; PRIOR FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: US 60/333, 545
; PRIOR FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: US 60/348, 464
; PRIOR FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/348, 615
; PRIOR FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/390, 804
; PRIOR FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: US 60/397, 557
; PRIOR FILING DATE: 2002-07-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1371
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: promemapsin 2-T1
; US-10-281-092-7

Alignment Scores:
Pred. No.: 7 42e-269 Length: 1371
Score: 2267.00 Matches: 429
Percent Similarity: 99.31% Conservative: 1
Best Local Similarity: 99.08% Mismatches: 3
Query Match: 99.08% Indels: 0
DB: 19 Gaps: 0

US-10-726-967a-84 (1-433) x US-10-281-092-7 (1-1371)
QY 1 THRGHNGIYTLIETRGLEUPLLEUARGSERGLYGLYALAPROLEUGLYLEU
DB 70 ACCCGACGAGGCGATCCGGCTGCCCTCGCGACGCGGCTGGGGGCGCCCTGGGGCTG 129
QY 21 GUUILEASNUGLUHTHRAAPGLUGLUPROGLUGLUPROGLYARGARGLYSERPHEVAI
DB 130 CGGCTGCCCGGAGAGCCGACGAGAGCCCGGAGAGCCCGGCGGAGGCGGCGAGCTTTGTG 189
QY 41 GIUMETVALAPASNULEUARGLYLYSERGLYGLNGLYTYRTRYVALGIUMETHTVAI
DB 190 GAGAGTGTGACACCTGAGGGGCAAGTGGGGCAGGGCTACTAGTGAAGTACCGCTG 249
QY 61 GLYSERPROPGINTHLEUASNULEUVALAPTHNGLYSERSEASNPHEALAVAI
DB 250 GGCACACCCCGGACGCTCAACATCTGGTGTGATACAGACGACGATTAACCTTGACGTG 309
QY 81 GLYLAALAPROHIEPROHIEUHIAPRTYTRYGNARGINLEUSERSETHTYR 100
DB 310 GGTGTGCCCCCGACCCCTTCTGTGATGTCTACTACCAAGAGCGTGTCCAGACATAC 369
QY 101 ARGAPLEUARGLYVALTYRVALPROTYRTHRGNGLYLSTPGLUGLYLEU 120
DB 370 CGGACCTCCGGAAGGTGTGTATGTGCTTACCCAGGCGCAAGTGGGAAGGGAGCTG 429
QY 121 GLYTHRAPLEUVALSERILEPROHISGLYPROASNAVITHRVALAASNIIEALA 140
DB 430 GGCACGACCTGTGATGATCCCCCAATGGCCCAACGTCATGTGCGTCCCAACATTGCT 489
QY 141 AIAIETHRGUSERASPLYSPHEPHEILEASNGLYSERASNTPTGULUGLYILEUGLY 160

```

```

DB 490 GCCATCATGATGACAGAAATTCTTCATCAACGGCTCCACCTGGAAAGCATCTGGCG 549
QY 161 LEUAIATYRALAGLILALARGPROASPHEPHEILEUPLUPROHIEPHEASSETLEU 180
DB 550 CTGGCTTGTCTGAATTTGCCAGGCTTGACATCTCCCTGGAGGCTTTCTTGTGACTCTTG 609
QY 181 VALLYSGINTHRHISVALPROASNULEUPHESELEUGINLEUCYSGLYALAGLYPHEPRO 200
DB 610 GTAAGCAGACCCAGCTTCCCACTCTTCTCCCTGACGCTTTGTGTGCTGGCTTCCC 669
QY 201 LEUANGINSEGLIVALLEUALASERYALGLYSERMETILEGLYGLYIIEASP 220
DB 670 CTCAACCGATGTGAAGTGTCTGCTGTGAGGAGCATGATCATTTGAGGATATGAC 729
QY 221 HISSELEUTYRTHRGLYSERLEUTRYTHRPROILEARGAGLUPTPTYTRYGLU 240
DB 730 CACTGCTGTACACAGGCGCTCTGGTATACCCATCCGCGGAGGTGATATTAG 789
QY 241 VALLIETLEVALARGVALGLUIEASNGLYGLNASPLEULYSWEAPCYSLYGLUTYR 260
DB 790 GTCATCATTTGCGCGGTGAGATCAATGACAGGATCTGAATAATGACCTGCACAGAGTAC 849
QY 261 ASNTYRASPYSERILEVALASPSERYLYTHRTHRASNULEUARGLEUPROLYSVAL 280
DB 850 AACTATGACAGACATTTGAGACATGAGCAACCACTTTGTTGCTTGCACAGAAAGTG 909
QY 281 PHEGLUALAVALYLSERILEYALALASERSETHRGLUSYSPHEPROASPLY 300
DB 910 TTGAACTGTGACGCAATCATCATCAAGGCGCTTCCCTCCAGAGAAAGTTCCCTGATGT 969
QY 301 PHETRPLEUGLYGLUGINLEUVALCYETTPGLNALGLYTHRTHRPROTPASNIIEPHE 320
DB 970 TTCTGCTAGAGAGCGAGTGTGTGCTGCGAAGCAGCACCCCTTGGAAATTTTC 1029
QY 321 PROVALIIESEERLEUTYRLEUMETGLYLUVALTHRASNULEUPHEARGIETHTRI 340
DB 1030 CAGTGTCTCATCTCACTACCTTAATGGGTGAGGTTACCAACGATCTTCCGATCACCATC 1089
QY 341 LEUPROGLINGINTYRLEUARGPROVALGIUASPVALALATHSERGINASPSAPCYTRY 360
DB 1090 CTTCCGACGAAATATCTGCGGCGGACGAGTGTGCGCAACGCTTCCCAACGACCTTTAC 1149
QY 361 LYSPEHALIIESEERINSESERSETHRGLYTHRVALMERGLYALVALIIEMETGLUGLY 380
DB 1150 AAGTTGCCATCTCAACATCATCATCAGGCGACTGTATGGAGCTGTATTCATGAGAGGC 1209
QY 381 PHETYRVALVALPHEASPARGLAARGLYSARGILEGLYPHEALAVASERIALACYSHIS 400
DB 1210 TTCTACGTTGTCTTTGATCGGGCCGAAAACGAAATTGGCTTGTCTGACGCCCTTGCCAT 1269
QY 401 VALHISAPGLUPHEARGTHRALAVALAGLYGLYPROHIEVALTHREUASPMETGLU 420
DB 1270 GTCCAGATGAGTTCAGAGCGGCGGAGGAGGCGCTTTGTGACCTTGACATGAGAA 1329
QY 421 ASPCYSGLYTYRASNIIEPROGLINTHRASPGUSERTHR 433
DB 1330 GACTGTGGCTTACACATTTCCACAGACAGATGAGTCAACC 1368

RESULT 2
US-11-089-918-1
; Sequence 1, Application US/11089918
; Publication No. US20050164327A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, John P.
; APPLICANT: Bassi, Gurigdal
; APPLICANT: Doane, Minh Tam
; APPLICANT: Frigon, Normand
; APPLICANT: John, Varghese
; APPLICANT: Power, Michael
; APPLICANT: Simha, Sukanto
; APPLICANT: Tatsuono, Gwen
; APPLICANT: Tung, Jay

```

```

; APPLICANT: Wang, Shuwen
; TITLE OF INVENTION: Beta-Secretase Enzyme Compositions and
; FILE REFERENCE: 228-US-NEW2C9
; CURRENT APPLICATION NUMBER: US/11/089,918
; PRIOR FILING DATE: 2005-03-25
; PRIOR APPLICATION NUMBER: US 09/723,722
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 09/501,708
; PRIOR FILING DATE: 2000-02-10
; PRIOR APPLICATION NUMBER: US 09/471,669
; PRIOR FILING DATE: 1999-12-24
; PRIOR APPLICATION NUMBER: 60/119,571
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: 60/139,172
; PRIOR FILING DATE: 1999-06-15
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1503
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-11-089-918-1

Alignment Scores:
Pred. No.:      8,54e-269      Length:      1503
Score:          2267.00      Matches:      429
Percent Similarity: 99.31%      Conservative: 1
Best Local Similarity: 99.08%      Mismatches: 3
Query Match:      24          Indels:      0
DB:              Gaps:      0

US-10-726-967a-84 (1-433) x US-11-089-918-1 (1-1503)

QY      1  ThrGlnHiGly1YleArgLeuProlLeuAArgSerGly1LeuG1G1ValAlaProLeuGlyLeu 20
DB      64  ACCCAACAGCGGCTATCCGGCTGCGCCCTGCGAGCGGCTGGGGGCCCCCTGGGGGCTG 123
QY      21  Glu1LeuAsnLeuGluThraPrgLugLupProG1LupProG1YArgArgGlySerPheVal 40
DB      124  CGGCTCCCCGGGAGACCGACAGAGAGCCCGAGAGCCCGCGAGGGCAGCTTTG 183
QY      41  GluMetValAspAsnLeuArgGlyYsSerGlyGlnGlyTyTyTyValGluMetThrVal 60
DB      184  GAGATGTGGAACAACCTGAGGGGCAAGTCCGGGCAAGGCTACTACGTGAGATGACCGTG 243
QY      61  GlySerProProGlnThraLeuAn1LeuValAspThrGlySerSerAnPheAlaVal 80
DB      244  GGCAGCCCCCGCAGACGCTCAACATCTGTGTGATACAGGCAAGTAACTTTGACATG 303
QY      81  GlValAlaAlaProHisProPheLeuH1sArgTyTyGlnArgGlnLeuSerSerThrTy 100
DB      304  GGTGTGCCCCCACCCTCTCTCTGCTGCTACTACAGAGGAGAGCTGTCCACACATAC 363
QY      101  ArgAspLeuArgGlyGlyValTyTyValProTyThrGlnGlyYsTyTrpGluGluLeu 120
DB      364  CGGAGCTCCGGAAGGCTGTGTATGTGCTTACACCCAGGCAAGTGGAGGAGGAGCTG 423
QY      121  GlyThraPheLeuValSer1LePProHisGlyProAsnValThraValArgAlaAn1Leu 140
DB      424  GGCACCGAAGCTGTGTAAGCATCCCCCATGAGGCCCAACCTCACTGTGCTGCAACATTC 483
QY      141  Ala1LePThrGluSerAspAspPhePhe1LeAsnGlySerAnTrpGluGly1LeuGly 160
DB      484  GCCATCACTGAATCAAGACAGTCTTATCAACGAGCTTCAACTGGAGAGGCACTTGGG 543
QY      161  LeuAlaTyra1Glu1Leu1AaArgProAspAspSerLeuGluProPhePheAspSerLeu 180
DB      544  CTGGCGATGCTGAGATGCGACAGGCTGACGAGCTCCCTGAGGCTTTCTTGAAGCTCTG 603
QY      181  Val1LeGlnThraHisValProAsnLeuPheSerLeuGlnLeuGlyGly1AaGlyPhePro 200
```

```

DB      604  GTAAGCAGACCACGTTCCCAACCTTCTCCCTGACGCTTTGTGTGCTGCTTCC 663
QY      201  LeuAsnGlnSerGluValLeuAlaSerValGlyGlySerMet1Le1Gly1LeuAsp 220
DB      664  CTCACACAGCTGAGAGCTGCTGCTGTGAGGAGACAGATCATTTGAGAGATTCAC 723
QY      221  HisSerLeuTyThrGlySerLeuTrpTyThrPro1LeaArgArgLupTyTyTyGlu 240
DB      724  CACTGCTGTACAGAGAGCTCTCGTATACACCCATCCGGCGGAGATGATTAATAG 783
QY      241  Val1Le1LeuValArgValGlu1LeAsnGlyGlnAspLeuTyMetAspCysLeuGlyTy 260
DB      784  GTGATCATTTGGCGGGTGGAGATCAATGACAGGATCTGAATAATGACTGCAGAGATAC 843
QY      261  AsnTyraAspGlySer1LeuValAspSerGlyThrThraAsnLeuArgLeuPro1LeuVal 280
DB      844  AACTATGACAGAGATGTGTGACAGTGGACACCAACCTTGCTTCCCAAGAAAGTG 903
QY      281  PheGluAla1AlaVal1LeSer1LeuVal1Ala1LeSerSerThraGluTyPheProAspGly 300
DB      904  TTGAAAGCTGCACTAAATTCATCAAGGCAAGCTCTTCCACGAGAAATTCCTGATG 963
QY      301  PheTrpLeuGlyGluGlnLeuValCysTrpGlnAlaGlyThraThrProTrpAsn1LePhe 320
DB      964  TTCTGGCTAGAGAGACAGCTGGTGTGTGACAGCAGACACCACTTGGAAATTTTC 1023
QY      321  ProVal1LeSerLeuTyLeuMetGlyGlyVal1ThraAsnGlnSerPheArg1LeThr1Le 340
DB      1024  CCAGTCATCTCACTCTACTATATGGGTGAGGTATACCAACCACTCTTCCGATCACATC 1083
QY      341  LeuProGlnGlnTyThrLeuArgProValGluAspVal1AlaThraSerGlnAspAspCysTy 360
DB      1084  CTTCCGAGCAATATCTTGCGGCAAGTGTGACAGTGTGCACTCCCAAGACAGACTGTTC 1143
QY      361  TyPheAla1LeSerGlnSerSerThraGlyThra1MetGlyVal1LeuMetGlyGly 380
DB      1144  AAGTTTGCATCTCAAGATTCACAGGCACTGTATATGGAGCTGTATATCATGAGGGGC 1203
QY      381  PheTyTyValValPheAspArgAlaArgYsArg1LeGlyPheAlaValSer1AaCysHis 400
DB      1204  TTCATGCTGTCTTGTATCTGGGCGCCGAAACCAATGGCTTTGCTGTCAGCCCTTGCCAT 1263
QY      401  Val1HisAspGluPheAspThrAla1AlaValGluGlyProPheValThraLeuAspMetGlu 420
DB      1264  GTGCAGATGATGATTCAGAGGCGAGCGGTGAGAGGCCCTTTGTCACTTGGACATGGA 1323
QY      421  AspCysGlyTyraSn1LeProGlnThraPrgLupSerThr 433
DB      1324  GACTGTGCTACACATTCACAGACAGATGATCAACC 1362

RESULT 3
US-11-090-866-1
; Sequence 1, Application US/11090866
; Publication No. US20050164294A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, John P.
; APPLICANT: Baal, Guribai
; APPLICANT: Doane, Minh Tam
; APPLICANT: Frigon, Normand
; APPLICANT: John, Varghese
; APPLICANT: Power, Michael
; APPLICANT: Simha, Sukanto
; APPLICANT: Tatsuno, Gwen
; APPLICANT: Tung, Jay
; APPLICANT: Wang, Shuwen
; APPLICANT: McConlogue, Lisa
; TITLE OF INVENTION: Beta-Secretase Enzyme Compositions and
; FILE REFERENCE: 228-US-NEW2C12
; CURRENT APPLICATION NUMBER: US/11/090,866
; CURRENT FILING DATE: 2005-03-25
; PRIOR APPLICATION NUMBER: US 09/723,722
; PRIOR FILING DATE: 2000-11-28
```

```

; PRIOR APPLICATION NUMBER: US 09/501,708
; PRIOR FILING DATE: 2000-02-10
; PRIOR APPLICATION NUMBER: US 09/471,669
; PRIOR FILING DATE: 1999-12-24
; PRIOR APPLICATION NUMBER: 60/119,571
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: 60/139,172
; PRIOR FILING DATE: 1999-06-15
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1503
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-11-090-866-1

Alignment Scores:
Pred. No.:      8,54e-269      Length:      1503
Score:          2267.00      Matches:      429
Percent Similarity: 99.31%      Conservative: 1
Best Local Similarity: 99.08%      Mismatches: 3
Query Match:      99.08%      Indels:      0
DB:               24          Gaps:         0

US-10-726-967a-84 (1-433) x US-11-090-866-1 (1-1503)

QY      1  ThrGlnHieGlyIleatgLeuProLeuAArgSerGlyLeuGlyGlyValAProLeuGlyLeu 20
Db      64  ACCGAGCAGCGGCAATCCGGCTGCCCTCGCAGCGCGCTGGGGGGGGCGCCCTGGGGCTG 123
QY      21  GluIleasLeuGluThrAepGluGluProGluGluProGlyAArgGlySerPheVal 40
Db      124  CGGCTGCCCGGAGAGACCGAGAGAGCCGAGAGAGCCGGCGGAGGGGCGAGCTTTGTG 183
QY      41  GluMetValAspAsnLeuAArgGlyLySerGlyGlyGlyThrValGluMetThrVal 60
Db      184  GAGATGATGACAACTCGAGGGGAGAGTGGGGCAGGGCTACTACGTGAGATGACCGTG 243
QY      61  GlySerProProGlnThrLeuAsnIleLeuValAspThrGlySerSerAsnPheAlaVal 80
Db      244  GGCAGCGCCCGCAGACGGCTCAACATCCGTGGATACAGGAGCAGTAACTTTGCAGTG 303
QY      81  GlyAlaAlaProHisProPheLeuHisArgIleValGlyArgGlyLeuSerSerThrTyr 100
Db      304  GGGTGTGCCCCCAGCCCTTCTCTGATCGCTACTACCAAGGAGAGCTGTCCAGCAGTAC 363
QY      101  ArgAspLeuAArgGlyValTyrValProTyrThrGlnGlyLySerThrGlnGlyLeu 120
Db      364  CGGAGCTCCGGAGGGGTGTATGTGCTTACACCCAGGGGCAAGTGGAGAGGGAGCTG 423
QY      121  GlyThrAspLeuValSerIleProHisGlyProAsnValThrValArgAlaAsnIleAla 140
Db      424  GGCACCGCACTGTAAAGCATCCCAATGCCCCCAAGTCACTGTGCGCAACTTGTCT 483
QY      141  AlaIleThrGlySerAspAspPhePheIleAsnGlySerAsnThrGlnGlyIleGly 160
Db      484  GCCATCATCGAATCGACAAAGTTCTTCAACAGGCTTCAACTGGAGAGGATCCTGGGG 543
QY      161  LeuAlaTyrAlaGluIleAlaArgProAspAspSerLeuGluProPhePheAspSerLeu 180
Db      544  CTGGCTATGCTGAGATTGCCAGGCTGACGACTCCCTGGAGGCTTTTCTTTGACTCTCTG 603
QY      181  ValIlyGlnThrHisValAProAsnLeuPheSerLeuGlnLeuCyGlyValAglYpHePro 200
Db      604  GTAAGAGAGACCCAGGTTCCCAACTCTTCTCCCTGCGAGCTTTGTGGTGGCTTCCCC 663
QY      201  LeuAsnGlnSerGluValIleuAlaSerValAglYpHeSerMetIleIleGlyIleAsp 220
Db      664  CTCACCAAGTCTGAAGTCTGCTCTCTCTGAGAGGAGACATGATGAGGATGAC 723
QY      221  HisSerLeuTyrThrGlySerLeuTyrThrProIleAArgGluTyrTyrGlu 240
Db      724  CACTCGCTGTACAGAGGAGTCTCTGTATACACCCCACTCCGGCGGAGTGTATTATGAG 783
```

```

QY      241  ValIleIleValAArgValGluIleAsnGlyGlnAspLeuLySerMetAspCyValGluTyr 260
Db      784  GTGATCATTTGGCGGGTGGAGATCAATGAGACGATCTGAAAATGGACTGCAGAGATAC 843
QY      261  AsnTyrAspLySerIleValAspSerGlyThrThrAsnLeuAArgLeuProLyVal 280
Db      844  AACTATGACAAAGAGATGTGACAGTGGCACCAACCTTCTGTGCCCAAGAAAGTG 903
QY      281  PheGluAlaAlaValLySerIleValAlaIleSerThrGlyLySerPheProAspGly 300
Db      904  TTTGAAGCTGACATCAATTCATCAAGCAGCTTCTCCACGAGAAAGTTCCTGATGT 963
QY      301  PheTyrLeuGlyGluGlnLeuValCysThrPheIleAlaGlyThrThrProTyrAsnIlePhe 320
Db      964  TTCTGGCTAGAGAGAGAGCTGTGTGGTGGCAACAGGACACCCCTTGGACATTTTC 1023
QY      321  ProValIleSerLeuTyrLeuMetGlyGluValThrAsnGlnSerPheAArgIleThrIle 340
Db      1024  CGAGTCATCTACCTTACCTTAATGGTGAAGTTACCAACAGTCTTCCGATCACCATC 1083
QY      341  LeuProGlnGlnTyrLeuAArgProValGluAspValAlaThrSerGlnAspAspCyTyr 360
Db      1084  CTTCCGAGAGCAATACCTGCGGCCAGTGGAGATGTGGCCAGTCCCAAGCGACTGTAC 1143
QY      361  LysPheAlaIleSerGlnSerSerThrGlyThrValMetGlyValAlaIleMetGluGly 380
Db      1144  AAGTTGCACTTCAACAGTATCCACCGGCACTGTATGGAGCTGTATCATGAGAGGGC 1203
QY      381  PheTyrValValPheAspAArgAlaArgLyValIleGlyPheAlaValSerAlaCyHis 400
Db      1204  TTCTACCTTGTCTTTATTCGGGCCCAAAACGAATGTGCTTGTCTGACGGCTTGCCAT 1263
QY      401  ValHisAspGluPheArgThrAlaAlaValGluGlyProPheValThrLeuAspMetGlu 420
Db      1264  GTGCACGATGAGTTACAGAGCGGAGCGGTGAAGGCCCTTTGTGTACCTTTGGACATGGA 1323
QY      421  AspCyGlyTyrThrAsnIleProGlnThrAspGluSerThr 433
Db      1324  GACTGTGGCTTAAACATTCACAGACAGATGATCAACC 1362

RESULT 4
; US-11-069-377-1
; Sequence 1, Application US/11069377
; Publication No. US20050170489A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, John P.
; APPLICANT: Basi, Guridpal
; APPLICANT: Doane, Minh Tam
; APPLICANT: Frigon, Normand
; APPLICANT: John, Varghese
; APPLICANT: Power, Michael
; APPLICANT: Simha, Sukanco
; APPLICANT: Tatsuno, Gwen
; APPLICANT: Tung, Jay
; APPLICANT: Wang, Shuwen
; APPLICANT: McConlogue, Lisa
; TITLE OF INVENTION: Beta-Secretase Enzyme Compositions and
; FILE REFERENCE: 228-US-NEW208
; CURRENT APPLICATION NUMBER: US/11/069,377
; CURRENT FILING DATE: 2005-02-28
; PRIOR APPLICATION NUMBER: US 09/723,722
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 09/501,708
; PRIOR FILING DATE: 2000-02-10
; PRIOR APPLICATION NUMBER: 60/119,571
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: 60/139,172
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
```



LENGTH: 1503  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-11-069-377-1

## Alignment Scores:

	8,54e-269	Length: 1503
Pred. No.:	2267.00	Matches: 429
Score:	99.31%	Conservative: 1
Percent Similarity:	99.08%	Mismatches: 3
Best Local Similarity:	99.08%	Indels: 0
Query Match:	24	Gaps: 0

US-10-726-967a-84 (1-433) x US-11-069-377-1 (1-1503)

```

QY      1 ThrGlnHsAGLYlLeArgLeuProLeuArgSerGlyLeuGlyGlyAlaProLeuGlyLeu 20
DB      64 ACCGACGACGCGATCCGCTCCCTGCGCAGCGGCTGGGGGCGCCCTCGGGGCTG 123
QY      21 GtUleAsnLeuGluThrAspGluGluProGluGluProGlyArgArgGlySerPheVal 40
DB      124 CGGCTGCGCCGGGAGACCGACGAGAGCGCGAGAGCGCGCGGAGGGGCGAGCTTTGCG 183
QY      41 GluMetValAspAsnLeuArgGlyLeuSerGlyGlyGlyGlyTyrTyrValGluMetThrVal 60
DB      184 GAGATGCTGACCACTGAGGGGCGACGCGGGCGAGGCTTACTTACGTGAGATGACCGTG 243
QY      61 GlySerProGluThrLeuAsnLeuValAspThrGlySerSerAsnPheAlaVal 80
DB      244 GCGAGCGCCCGGACGCTCAACATCTCGTGAGATACAGGACGACGATTACTTGCAGTG 303
QY      81 GlyAlaAlaProHsAspPheLeuHsArgTyrTyrGlnArgGlnLeuSerSerThrTyr 100
DB      304 GGGTGGCGCCCGACCCCTTCCGTCGATCGCTACTACAGAGGCGAGCTGTCCAGACATAC 363
QY      101 ArgAspLeuArgGlyGlyValTyrValProTyrThrGlnGlyGlyGlyGlyGlyGlyGly 120
DB      364 CGGAGCTCCGGAGGGGTGTATGTGCGCTTACACCGAGGCGAAGGGGAGGCTG 423
QY      121 GlyThrAspLeuValSerLeuProHsArgTyrProAsnValThrValArgAlaAsnIleAla 140
DB      424 GGCACCGACCTGGTACATCCCGCCCGCCCGACGCTACTGCTGCTGCCAATTCCT 483
QY      141 AlaIleThrGluSerAspLeuPhePheIleAsnGlySerAsnTrpGluGlyIleLeuGly 160
DB      484 GCCATCACTGAACACAGCAAGTCTTCACTCAACGCGCTCCACAGGAGGACATCTG 543
QY      161 LeuAlaTyrAlaGluIleAlaArgProAspAspSerLeuGluProPhePheAspSerLeu 180
DB      544 CTGGCTTATGCTGAGATTGCGCAGGCTGACGACTCCCTGAGGCTTCTTGACTCTCTG 603
QY      181 ValIleGluThrHisValProAsnLeuPheSerLeuGlnLeuGlyGlyAlaGlyPhePro 200
DB      604 GTAAACAGACCCACGCTCCCAACCTTCTCCCTGAGCTTGTGGTCTGCTTCC 663
QY      201 LeuAsnGlnSerGluValLeuAlaSerValGlyGlySerMetIleIleGlyGlyIleAsp 220
DB      664 CTCACACAGCTGAAGTGGCTGCTGCTGCGAGGGAGCATGATCATTTGAGATAC 723
QY      221 HisSerLeuTyrThrGlySerLeuTyrTyrThrProIleArgArgGluTyrTyrGlu 240
DB      724 CACTCGCTGTACACAGGACAGTCTCTGTGTATACACCCATCGGGGGAGGTGTATTGAG 783
QY      241 ValIleIleValArgValGluIleAsnGlyGlnAspLeuMetAspCysIleGluTyr 260
DB      784 GTATCATTTGTGGGGGAGATCATATGACAGGATTTGAAAATGACATCAGAGATAC 843
QY      261 AsnTyrAspLeuSerIleValAspSerGlyThrThrAsnLeuArgLeuProGlyIleVal 280
DB      844 AACTATGACAGAGCATTTGTGACAGTGGCACCACTTGTGTTGGCCCAAGAAAGTG 903
QY      281 PheGluAlaAlaValIleYsSerIleValAlaSerSerThrGluYsPheProAspGly 300

```

```

DB      904 TTTGAGCTGACGATCAATTCATCAAGGACGCTCTCCACGAGAGAGTTCCTGATGAT 963
QY      301 PheTrpLeuGlyGluGlnLeuValCysTrpGlnAlaGlyThrThrProTrpAsnIlePhe 320
DB      964 TTCTGGCTAGGAGGACACTGGTGTGTGCGGACAGGACACACCCCTTGAACATTTTC 1023
QY      321 ProValIleSerLeuTyrLeuMetGlyGlyValValThrAsnGlnSerPheArgIleThrIle 340
DB      1024 CAGTCATCTCACTTACTTAATGGGTGATGATACCAACAGTCTCTTCGATACCAATC 1083
QY      341 LeuProGlnGlnTyrLeuArgProValGluAspValAlaThrSerGlnAspAspCysTyr 360
DB      1084 CTTCCGACGACATACCTCGCGGACAGTGAAGATGTGGCCAGTCCCAAGACGACTGTAC 1143
QY      361 LysPheAlaIleSerGlnSerSerThrGlyThrValMetGlyAlaValIleMetGlyGly 380
DB      1144 AAGTTTGCATCTCCACAGTATCCACGGGCACTGTATAGGAGCTGTATATCATGAGGGC 1203
QY      381 PheTyrValAlaPheAspArgAlaArgLysArgIleGlyPheAlaValSerAlaCysHis 400
DB      1204 TTCTACGTGTCTTGTATCGGCGCCGAAAGCAATGGCTTTGCTGTACGCCCTTGCCAT 1263
QY      401 ValHisAspGluPheArgThrAlaAlaValGluGlyProPheValThrLeuAspMetGlu 420
DB      1264 GTGCACGATGATTCAGAGCGGCAAGCGGTGAGAGGCCCTTTGTCACTTGACATGMA 1323
QY      421 AspCysGlyTyrAsnIleProGlnThrAspGluSerThr 433
DB      1324 GACTGTGGCTACACATTCACAGACAGATGATGATCAAC 1362

```

## RESULT 5

```

US-11-090-872-1
; Sequence 1, Application US/11090872
; Publication No. US2005017788A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, John P.
; APPLICANT: Bassi, Gurigbal
; APPLICANT: Doane, Minh Tam
; APPLICANT: Frigon, Normand
; APPLICANT: John, Varghese
; APPLICANT: Power, Michael
; APPLICANT: Sinha, Sukanto
; APPLICANT: Tatsuno, Gwen
; APPLICANT: Tung, Jay
; APPLICANT: Wang, Shuwen
; APPLICANT: McConlogue, Lisa
; TITLE OF INVENTION: Beta-Secretase Enzyme Compositions and
; TITLE OF INVENTION: Methods
; FILE REFERENCE: 228-US-NEW2C11
; CURRENT APPLICATION NUMBER: US/11/090,872
; PRIOR FILING DATE: 2005-03-25
; PRIOR APPLICATION NUMBER: US 09/723,722
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 09/501,708
; PRIOR FILING DATE: 2000-02-10
; PRIOR APPLICATION NUMBER: US 09/471,669
; PRIOR FILING DATE: 1999-12-24
; PRIOR APPLICATION NUMBER: 60/119,571
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: 60/139,172
; PRIOR FILING DATE: 1999-06-15
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1503
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-11-090-872-1

```

## Alignment Scores:

	8,54e-269	Length: 1503
Pred. No.:	2267.00	Matches: 429
Score:	99.31%	Conservative: 1
Percent Similarity:	99.08%	

Best Local Similarity: 99.08% Mismatches: 3  
Query Match: 99.08% Indels: 0  
DB: 24 Gaps: 0

US-10-726-967a-84 (1-433) x US-11-090-872-1 (1-1503)

```
QY 1 ThGlnHieGlyIleAsgLeuProLeuAsgSerGlyLeuGlyValAProLeuGlyLeu 20
DB 64 ACCGACGACGACATCCGGCTGCCCCGCGACGCGCTGGGGGGCGCCCCCTGGGGCTG 123
QY 21 GluIleAsnLeuGluThrAsgGluGluProGluGluProGlyAArgAsgSerPheVal 40
DB 124 CGGCTGCCCCGGGAGACGACGAAGGCCGAGAGGCCGCGCGAGGGGACACTTTGTC 183
QY 41 GluMetValAspAsnLeuAsgGlyLeuSerGlyGlnGlyTyrTyrValGluMetThrVal 60
DB 184 GAGATGGTGGACAACTCGAGGGGCAAGTCGGGGCAGGGCTACCTGCGAGATGACCGTG 243
QY 61 GlySerProGluInThrLeuAsnIleLeuValAAspThrGlySerSerAsnPheAlaVal 80
DB 244 GCGAGCCCCCGACGACGCTCAACATCTGGTGATACAGGACAGCAAGTAACTTGGACATG 303
QY 81 GlyAlaAlaProHisProPheLeuH1AsgTyrTyrGlnAsgGlnLeuSerThrTyr 100
DB 304 GGTGCTGCCCCCACCCTTCTCGCATCGCTACTACAGAGGACGCTGCCAGACATAC 363
QY 101 ArgAspLeuArgGlyGlyValTyrValProTyrThrGlnGlyLeuTyrPgluGlyLeu 120
DB 364 CGGACCTCCGGAAGGGTGTATGTGCCCCCTACACGAGGCAAGTGGAGAGGGAGCTG 423
QY 121 GlyThrAspLeuValSerIleProHisGlyProAsnValThrValArgAlaAsnIleAla 140
DB 424 GGCACGACCTGGTATAGCATCCGCCATGGCCCCCAAGTCACTGGGTGCCAATGGT 483
QY 141 AlaIleThrGluSerAspLeuPhePheIleAsnGlySerAsnTyrPgluGlyLeuGly 160
DB 484 GCCTATCAGTAATCAACAAGTCTTTCATCAACGCGCTCCACGCGGAGAGGACCTCGGG 543
QY 161 LeuAlaTyrAlaGluIleAlaArgProAspAspSerLeuGluProPhePheAspSerLeu 180
DB 544 CTGGCTATGCTGAGATTCGACGCTTCCAGGCTTGACACTCCCTGAGACCTTTCTTATACCTCG 603
QY 181 ValIleGlnThrHisValProAsnLeuPheSerLeuGlnLeuGlyAlaGlyPhePro 200
DB 604 GTAAAGCAGACCCAGCTTCCCAACCTCTCTCCCTGACGCTTGTGGGTGGCTGCC 663
QY 201 LeuAsnGlnSerGluValLeuAlaSerValGlyLeuSerMetIleIleGlyGlyLeuAsp 220
DB 664 CTCAACCAAGTCTGAAGTGTGGCTCTCTCGGAGGAGCATGATCATTTGGAGATTCGAC 723
QY 221 HisSerLeuTyrThrGlySerLeuTyrTyrThrProIleArgAsgGluTyrTyrGlu 240
DB 724 CACTCGCTGTACACAGGACGCTCTGTGTRTACACCACTCCGCGGAGGTGATTTATGAG 783
QY 241 ValIleIleValArgValGluIleAsnGlyIleAspLeuLeuMetAspCysLeuGlyLeuTyr 260
DB 784 GGTATCATTTGTCGGGTGAGATCAATGACACGATCTGAAAATGACTGCACAGGAGTAC 843
QY 261 AsnTyrAspLeuSerIleValAspSerGlyThrThrAsnLeuArgLeuProLeuValVal 280
DB 844 AACTATGACAGACATTTGTGACAGTGACACCAACCACTTCGTTTCCCAAGAAATG 903
QY 281 PheGluAlaAlaValLeuSerIleValAlaAspSerThrGluValPheProAspGly 300
DB 904 TTTGAAAGCTGAGCATCAATTCATCAAGGACGCTCTCCACGAGAAATTCCTGATGCT 963
QY 301 PheTyrPleuGlyGluGlnLeuValCysTyrPgluAlaGlyThrThrProTyrAsnIlePhe 320
DB 964 TTTGCGCTAGAGACACACTGGTGTGCTGGCAAGGACCAACCCCTTGGAAATTTTC 1023
QY 321 ProValIleSerLeuTyrLeuMetGlyGluValThrAsnGlnSerPheArgIleThrIle 340
DB 1024 CCAAGTCACTCACTTAACTTAATGGGTAGGTTACCAACCAAGTCTTCGATCAACATC 1083
```

```
QY 341 LeuProGlnGlnTyrLeuArgProValGluAspValAlaThrSerGlnAspAspCysTyr 360
DB 1084 CTTCGCGACAAATACCTCGGCGACGTGGAATGTGGCACACTCCAAAGACACTGTAC 1143
QY 361 LysPheAlaIleSerGlnSerSerThrGlyThrValMetGlyAlaValIleMetGluGly 380
DB 1144 AAGTTTGCATCTCACTCATTCACACGGGCTGTATGGAGCTGTATCATCATGAGAGGC 1203
QY 381 PheTyrValAlaPheAspArgAlaArgGlyValArgIleGlyPheAlaValSerAlaCysHis 400
DB 1204 TTTCACTGTTCTTTGATCGGCGCCGAAACGAATTTGGCTTGTCTGACGGCTTGGCAT 1263
QY 401 ValHisAspGluPheArgThrAlaAlaValGluGlyProPheValThrLeuAspMetGlu 420
DB 1264 GTGCACGATGATTCAGGACGCGAGCGGTGGAAAGCCCTTTTGTCACTTGGACATGGA 1323
QY 421 AspCysGlyTyrAsnIleProGlnThrAsgGluSerThr 433
DB 1324 GACTGTGCTACCAACTTCACAGACAGATGATCAACC 1362
```

## RESULT 6

```
US-10-214-932-103
; Sequence 103, Application US/10214932
; Publication No. US20030100707A1
; GENERAL INFORMATION:
; APPLICANT: HWANG, Inman
; APPLICANT: KIM, Dae Heon
; APPLICANT: LEE, Yong Jik
; TITLE OF INVENTION: SYSTEM FOR DETECTING PROTEASE
; FILE REFERENCE: APB02/US
; CURRENT APPLICATION NUMBER: US/10/214,932
; CURRENT FILING DATE: 2002-08-08
; NUMBER OF SEQ ID NOS: 133
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 103
; LENGTH: 1506
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: gene
; LOCATION: (1)..(1506)
; OTHER INFORMATION: Gene for APP beta-secretase
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1503)
; OTHER INFORMATION: APP beta-secretase
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank/AF201468
; DATABASE ENTRY DATE: 1999-12-19
US-10-214-932-103
```

## Alignment Scores:

```
Pred. No.: 8,57e-269 Length: 1506
Score: 2267.00 Matches: 429
Percent Similarity: 99.31% Conservative: 1
Best Local Similarity: 99.08% Mismatches: 3
Query Match: 99.08% Indels: 0
DB: 15 Gaps: 0
```

US-10-726-967a-84 (1-433) x US-10-214-932-103 (1-1506)

```
QY 1 ThGlnHieGlyIleAsgLeuProLeuAsgSerGlyLeuGlyValAProLeuGlyLeu 20
DB 64 ACCGACGACGACATCCGGCTGCCCCGCGACGCGCTGGGGGGCGCCCCCTGGGGCTG 123
QY 21 GluIleAsnLeuGluThrAsgGluGluProGluGluProGlyAArgAsgSerPheVal 40
DB 124 CGGCTGCCCCGGGAGACGACGAAGGCCGAGAGGCCGCGGAGGGGACGCTTTGTC 183
QY 41 GluMetValAspAsnLeuAsgGlyLeuSerGlyGlnGlyTyrTyrValGluMetThrVal 60
DB 184 GAGATGGTGGACAACTCGAGGGGCAAGTCCGGGGCAGGGCTACTAGTGAATGACCGTG 243
```

QY 61 GYserProGlnThrLeuasnileuValAspThrGlySerSeranphealVal 80  
 DB 244 GGCACGCCCCCGACAGCGCTCAACATCTGGTGAATACAGGACAGCATTTGACAGT 303  
 QY 81 GYAlaAlaProHisProPheleuHisArgTyrTyrGlnArgGlnLeuSerThrTyr 100  
 DB 304 GGTGTGCCCCCGACCCCTTCTGCACTGCTACTACAGGACGAGCTGTCAGCAGATAC 363  
 QY 101 ArgAspLeuArgGlyValTyrValProTyrThrGlnGlyLeuTyrPgluGluLeu 120  
 DB 364 CGGACCTCCGGAAGGAGTGTATGTGCTTACACCCAGGACAGTGGAGAGGAGCTG 423  
 QY 121 GYThrAspLeuValSerileProHisGlyProasnValThrValArgAlaasnileal 140  
 DB 424 GGCACGACCTGGTAAAGCATCCCGACGCCCCCAAGCATGTCGTGCTCAACATTTGCT 483  
 QY 141 AlaileThrGluSerAspLysPhePheileasnGlySerSeranThrGluGlyleuGly 160  
 DB 484 GCCATCACTGAATCAGACAAAGTTCTTCATCAAGGCTCCAACTGGAGAGGACATCTGG 543  
 QY 161 LeuAlaTyrAlaGluileAlaArgProAspAspSerleuGlnLeuGlyValaglyPhePro 180  
 DB 544 CTGGCTATGCTGAATGCGACGCTGACGACTCCCTGAGGCTTTCTTGACTCTG 603  
 QY 181 VallyleGlnThrHisValProasnLeuPheSerleuGlnLeuGlyValaglyPhePro 200  
 DB 604 GTPAAGCAACCCACCTTCCCAACCTTCTCTCCAGCTTTGTGTGCTGCTGCTCC 663  
 QY 201 LeuAsnGlnSerGluValleuAlaSerValGlyGlySerMetleileGlyValleuAsp 220  
 DB 664 CTCACACAGTGAAGTGTGCTGCTGCTGCGAGGAGAGATGATTCATGAGATTCAG 723  
 QY 221 HisSerleuTyrThrGlySerleuTyrTyrThrProGlnArgGlnTyrTyrTyrGly 240  
 DB 724 CACTGCTGTACACAGGAGCTCTCTGTATACCCATCCGCGGAGTGTATTTAG 783  
 QY 241 ValileileValArgValGluileasnGlyGlnAspLeuLysMetAspCysLysGluTyr 260  
 DB 784 GTCATCATTTGTGGGGTGGAGATCAATGACAGATCTGAATTAATGACTCAAGAGTAC 843  
 QY 261 AsnTyrAspLysSerileValAspSerGlyThrThrAsnLeuArgLeuProLysVal 280  
 DB 844 AACTATGACAAAGCATTTGTGACAGTGGACCAACCACTTGTGTTGCCAAGAAAGT 903  
 QY 281 PheGluAlaAlaValLysSerileLysAlaAlaSerSerThrGluLysPheProAspGly 300  
 DB 904 TTTGAAGCTGCACTCAATCATCAAGGACGCTCTCCACGAGAGATTCCTGATGCT 963  
 QY 301 PheTyrLeuGlyGluGlnLeuValCysTyrGlnAlaGlyThrThrProTyrAsnilePhe 320  
 DB 964 TTCTGCTAGAGAGAGAGCTGTGTGTGCTGCGAAGGACGACCCCTTGAAATTTTC 1023  
 QY 321 ProValileSerleuTyrLeuMetGlyGluValThrAsnGlnSerPheArgileThrile 340  
 DB 1024 CCAGTCACTCACTCACTCAATAGGTGAGTTACCAACAGATCTTCCCATCACCATC 1083  
 QY 341 LeuProGlnGlnTyrLeuArgProValGluAspValAlaThrSerGlnAspAspCysTyr 360  
 DB 1084 CTTCCGAGCAATACCTGCGGCGAGTGAAGATGTGCCAGCTCCCAAGCAAGCTTTAC 1143  
 QY 361 LysPheAlaileSerGlnSerSerThrGlyThrValMetGlyAlaValileMetGluGly 380  
 DB 1144 AAGTTTGCCATCTCAAGTATCCACGGGACCTGTTATAGGAGCTGTTATCATGAGAGGC 1203  
 QY 381 PheTyrValAlaPheAspArgAlaArgLysArgileGlyPheAlaValSerAlaCysHis 400  
 DB 1204 TTCTACGTTGCTTGTGATCGGGCGCGAAGAAAGATTTGTGCTGTCACGCTTGCAT 1263  
 QY 401 ValHisAsnGluPheArgThrAlaAlaValGluGlyProPheValThrLeuAspMetGlu 420  
 DB 1264 GTGCACGATGATTCAGAGAGGCGGCGGTGAAGGCGCTTTGTGACCTTGGACATGAA 1323

QY 421 AspCysGlyTyrAsnileProGlnThrAspGlnSerThr 433  
 DB 1324 GACTGTGCTACACATTCACAGACAGATGATGATACC 1362

RESULT 7  
 US-10-281-092-5  
 Sequence 5, Application US/10281092  
 Publication No. US20040121947A1  
 GENERAL INFORMATION:  
 APPLICANT: Ghosh, Arun K.  
 APPLICANT: Tang, Jordan J.N.  
 APPLICANT: Bilcer, Geoffrey  
 APPLICANT: Chang, Wampin  
 APPLICANT: Hong, Lin  
 APPLICANT: Koelsch, Gerald E.  
 APPLICANT: Loy, Jeffrey A.  
 APPLICANT: Turner, Robert T., III  
 APPLICANT: Devanadrum, Thippeswamy  
 TITLE OF INVENTION: COMPOUNDS WHICH INHIBIT BETA-SECRETASE  
 TITLE OF INVENTION: ACTIVITY AND METHODS OF USE THEREOF  
 FILE REFERENCE: 2932.1001-004  
 CURRENT APPLICATION NUMBER: US/10/281,092  
 PRIOR FILING DATE: 2002-10-23  
 PRIOR APPLICATION NUMBER: US 10/032,818  
 PRIOR FILING DATE: 2001-12-28  
 PRIOR APPLICATION NUMBER: PCT US01/50826  
 PRIOR FILING DATE: 2001-12-28  
 PRIOR APPLICATION NUMBER: US 60/258,705  
 PRIOR FILING DATE: 2000-12-28  
 PRIOR APPLICATION NUMBER: US 60/275,756  
 PRIOR FILING DATE: 2001-03-14  
 PRIOR APPLICATION NUMBER: US 60/335,952  
 PRIOR FILING DATE: 2001-10-23  
 PRIOR APPLICATION NUMBER: US 60/333,545  
 PRIOR FILING DATE: 2001-11-27  
 PRIOR APPLICATION NUMBER: US 60/348,464  
 PRIOR FILING DATE: 2002-01-14  
 PRIOR APPLICATION NUMBER: US 60/348,615  
 PRIOR FILING DATE: 2002-01-14  
 PRIOR APPLICATION NUMBER: US 60/390,804  
 PRIOR FILING DATE: 2002-06-20  
 PRIOR APPLICATION NUMBER: US 60/397,557  
 PRIOR FILING DATE: 2002-07-19  
 Remaining Prior Application data removed - See File Wrapper or PALM.  
 NUMBER OF SEQ ID NOS: 59  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 5  
 LENGTH: 1506  
 TYPE: DNA  
 ORGANISM: Unknown  
 FEATURE:  
 OTHER INFORMATION: memapsin 2  
 US-10-281-092-5

Alignment Scores:  
 Pred. No.: 8,576-269 Length: 1506  
 Score: 2267.00 Matches: 429  
 Percent Similarity: 99.31% Conservative: 1  
 Best Local Similarity: 99.08% Mismatches: 3  
 Query Match: 99.08% Indels: 0  
 DB: 19 Gaps: 0

US-10-726-967a-84 (1-433) x US-10-281-092-5 (1-1506)

QY 1 ThrGlnHisGlyileArgLeuProLeuArgSerGlyLeuGlyValAlaProLeuGlyLeu 20  
 DB 64 ACCCAGCAGCGGATCCGGCTGCTGCGCAGCGGCTGGGGGGCGCCCCCTGGGGCTG 123  
 QY 21 GlnileAsnLeuGlnThrAspGlnGluGluProGlnGluPgluArgAlaArgLysSerPheVal 40  
 DB 124 CGGCTGCCCGGAGAGCCGACGAAAGAGCCCGAGAGACCCGCGGAGGGGACATTTGTG 183  
 QY 41 GluMetValAspAsnLeuArgGlyLysSerGlyGlnGlyTyrTyrTyrValGluMetThrVal 60

Db	184	GAGATGGTGGCAACCTTGAGGGGCAAGTCGGGGGAGGGCTTACATCGAGAGATGACCGTG	243
Qy	61	GIYserProProGlnThrIleuAsnIIleuValAspThrGIYserSerAspPheAlaVal	80
Db	244	GGGAGCCCCCGGAGAGGCTCAATCTCTGGTGGATACAGGCAAGCAATACTTTGGCAGTG	303
Qy	81	GIYAlaIalProHisProPheIleuHisArgTYrTYrGlnArgGlnIleuSerSerThrTYr	100
Db	304	GGTGGTGGCCCCCAACCCCTTCCTGGCAATCGCTACACAGAGGCGAGCTGTCCAGACATAC	363
Qy	101	ArgAspLeuArgIYsGIYValTYrValProTYrThrGlnGlySerTPsIUGIYGIuLeu	120
Db	364	CGGGAGCTCCGGAAAGGCTGTGTATGTCCCTTACACCAGGAGGCAATGGAAAGGGAGCTG	423
Qy	121	GIYThrAspLeuValSerIIleProHisGlyProAsnValThrValArgAlaAsnIIleAla	140
Db	424	GGGACCGCACTGGTAAGCATCCCCCAAGGCCCAACCTCATCTGGCGCTGCCCAATTGCT	483
Qy	141	AlaIIleThrGlnSerAspIYsPhePheIIleAsnGlySerAsnTPsIUGIYIleuGly	160
Db	484	GGCATCACTAATCAACACAAATTCCTTCATCAAGGCTCCAACTGGGAAGGCATCTGGGG	543
Qy	161	LeuAlaTYrAlaGlnIleAlaArgProAspAspSerLeuGlnProPheAspSerLeu	180
Db	544	CTGGGCTATGCTGAGATTGGCAGGCTGACGATCCCTGGAGGCTTTCTTTGACTCTTG	603
Qy	181	ValIYsGlnThrHisValProAsnIleuPheSerLeuGlnIleuCYsgIYAlaGlyPhePro	200
Db	604	GTAAGAGAGACCCACGTTCCCAACCTTCCTCCATGGTGTGGTGGCTTCCCC	663
Qy	201	LeuAsnGlnSerGluValLeuAlaSerValGlyIYserMetIIleIleGlyIYIleAsp	220
Db	664	CTCAACAGCTGAAGTGGCTGGCTCTGTGGAGGAGACATGATCATTTGAGGATACGAC	723
Qy	221	HisSerLeuTYrThrGlySerLeuTYrTYrThrProIleArgArgGluTYrTYrGln	240
Db	724	CAGTCCCTGTACACAGCAGGTCTCTGTATACACCATCCGGGGGAGGTGATATATGAG	783
Qy	241	ValIIleIleValArgValGlnIleAsnGlyGlnAspLeuIleMetAspCYsIYsGIuTYr	260
Db	784	GTCATCATTTCTGGGGGTGGAGATCAATGACAGAGATCTGAANAATGCACTGCAGAGATAC	843
Qy	261	AsnTYrAspIYsSerIIleValAspSerGlyTYrThrAsnLeuArgLeuProIYsIYsVal	280
Db	844	AACATATGACAAAGCATTTGTGACAGTGGCACACCAACCTTCGTTGGCCAAAGAAAGTG	903
Qy	281	PheGlnAlaIleValIYsSerIIleYsAlaIleSerSerThrGlnIYsPheProAspGly	300
Db	904	TTTGAAGCTGCAGTCAAAATTCATCAAGGAGCTCTCCACGAGAAAGTTCCCTTAAGGT	963
Qy	301	PheTYrLeuGlyGlnGlnIleuValCYsTPsGlnAlaGlyTYrThrProTYrAsnIIlePhe	320
Db	964	TTCTGGCTAGGAGAGAGAGCTGGTGTGCTGGCAAGGACGACACCCCTTGGAAACTTTTC	1022
Qy	321	ProValIIleSerLeuTYrIleuMetGlyGluValThrAsnGlnSerPheArgIIleThrIle	340
Db	1024	CCAGTCATCTCACTCACTAATATGGGTGGGTATACCAACAAATCTTCGCACTCACATC	1083
Qy	341	LeuProGlnGlnTYrLeuArgProValGlnAspValAlaThrSerGlnAspAspCYsTYr	360
Db	1084	CTTCCGACGACAAATACCTGGGGCCAGTGAAGATGTGGCCAGCTCCCAAGACGACTGTTC	1143
Qy	361	IYsPheAlaIIleSerGlnSerSerThrGlyTYrValIleMetGlyAlaValIIleMetGlnIY	380
Db	1144	AAGTTGGCATCTCCAGTCATCCACGGGCACTGTTATGGGAGCTGTATCATGAGGGGC	1203
Qy	381	PheTYrValIleAlaPheAspArgAlaArgIYsArgIIleGlyPheAlaValSerAlaCYsHis	400
Db	1204	TTCTACGTTGCTTTGATCGGGCCGAAACGAATTTGGCTTGTGCTGACGGCTTGGCAT	1263
Qy	401	ValHisAspGluPheArgTYrAlaIleAlaValGlnGlyProPheValThrLeuAspMetGlu	420

Db	1264	GGTCACGATGAGTTCCAGACGGCAGCGGTGGAAGGCCCTTTTGTCACTTGACCTTGACATGAA	1322
Qy	421	APPVAGLTYTyrAsnLLeProGlnThrAspGluSerThr	433
Db	1324	GACTGTGCTTACAACTTCACAGACAGATGATGACACC	1362
RESULT 8			
US-10-275-339A-6			
; Sequence 6, Application US/10275339A			
; Publication No. US20040110743A1			
; GENERAL INFORMATION:			
; APPLICANT: MIYAMOTO, Masaomi			
; APPLICANT: MATSUI, Junji			
; APPLICANT: FUKUMOTO, Hiroaki			
; APPLICANT: TARUI, Naoki			
; TITLE OF INVENTION: Beta Secretase Inhibitors			
; FILE REFERENCE: 2729 USOP			
; CURRENT APPLICATION NUMBER: US/10/275,339A			
; CURRENT FILING DATE: 2003-10-30			
; PRIOR APPLICATION NUMBER: PCT/JP01/04144			
; PRIOR FILING DATE: 2001-05-18			
; PRIOR APPLICATION NUMBER: JP 2000-152758			
; PRIOR FILING DATE: 2000-05-19			
; NUMBER OF SEQ ID NOS: 9			
; SOFTWARE: PatentIn version 3.2			
; SEQ ID NO 6			
; LENGTH: 1527			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-10-275-339A-6			
Alignment Scores:			
Pred. No.: 8,75e-269 Length: 1527			
Score: 2267.00 Matches: 429			
Percent Similarity: 99.31% Conservative: 1			
Best Local Similarity: 99.08% Mismatches: 3			
Query Match: 99.08% Indels: 0			
DB: 19 Gaps: 0			
US-10-726-967A-84 (1-433) x US-10-275-339A-6 (1-1527)			
Qy	1	ThrglnHsGlyIleArgLeuProLeuArgSerGlyLeuGlyAlaProLeuGlyLeu	20
Db	64	ACCAGCAGCAGGATCCGGCTGCGCTCGCGCAGCGCGCTGGGGGGCGCCCCCTGGGGCTG	123
Qy	21	GluLeuAsnLeuGluThrAspGluGluProGluGluProGlyArgArgGlySerPheVal	40
Db	124	CGGCTGCCCCGGGAGACCGAGAAAGAACCCGAGGAGCCCGCGCGAGGGGCACTTTGTG	183
Qy	41	GluMetValAspAsnLeuArgGlyLysSerGlyGlnGlyTyrTyrValGluMetThrVal	60
Db	184	GAGATGGTGGACAACCTCGAGGGGCAAGCTCGGGGCAAGGCGCTACTGATGATGACCGTG	243
Qy	61	GlySerProProGlnThrLeuAsnIleLeuValAspThrGlySerSerAsnPheAlaVal	80
Db	244	GGCGAGCCCCCGCAGACGCTCAACATCTGTGGATACAGGCGACGTAACCTTTCAGCTG	303
Qy	81	GlyAlaAlaProHisProPheLeuHisArgTyrTyrGlnArgGlnLeuSerSerThrTyr	100
Db	304	GGTCTGCCCCCGCCCTCTCTGCAATCGTACTTACAGAGSAGAGCTGTCCAGACATAC	363
Qy	101	ArgAspLeuArgLysGlyValTyrValProTyrThrGlnGlyLysTyrGluGluLeu	120
Db	364	CGGAGACCTCCGGAAAGGCTGTGTATGTGCCCTTACACCGAGGGCAAGGGAGGAGCTG	423
Qy	121	GlyThrAspLeuValSerIleProHisGlyProAsnValThrValArgAlaAsnIleAla	140
Db	424	GGACCGAAGCTGGAGACATCCCGCAATGGCCCCCAAGCTCATGTGTCACCAATTGCT	483
Qy	141	AlaIleThrGluSerAspLysPhePheIleAsnGlySerAsnTrpGluGlyIleLeuGly	160
Db	484	GCCATTCATGAATCAAGCAAGTTCTTTCATCAACGGCTCCACCTGGGAAGCATCTCTGGGG	543

```
QY 161 LeuAlaTYrAlaGluIleAlaArgProAspAspSerLeuGluProPhePheSerLeu 180
DB 544 CTGGCTTAAGCTGAGATTCGACGGCTGACGACTCCCTGAGCCTTTCTTGACTCTCG 603
QY 181 ValIleValThrHisValProAsnLeuPheSerLeuGluLeuCyseGlyValaGlyPhePro 200
DB 604 GTAAGACAGACCACCTTCCCAACCTCTCTCCCTGACACTTGTGGTCTGGCTTCCCC 663
QY 201 LeuAsnGlnSerGluValLeuAlaSerValaGlyGlySerMetIleIleGlyGlyIleAsp 220
DB 664 CTCACACGACTGAGAGTGTGGCTCTGTGAGGAGGATGATCATGAGGATGAC 723
QY 221 HisSerLeuTYrThrGlySerLeuTYrTYrThrProIleArgArgGluTYrTYrGlu 240
DB 724 CACTCGCTTAACAGCAGGAGCTCTGTGTATACCCATCCGGGAGGTGATTTATGAG 783
QY 241 ValIleIleValArgValaGluIleAsnGlyGlnAspLeuLeuMetAspCysIleGlyIle 260
DB 784 GTGATCATTTGTGGGGTGGAGATCATGACAGATCTGAAAATGAGACTGCAAGAGTAC 843
QY 261 AsnTYrAspIleSerIleValaAspSerGlyTYrThrAsnLeuArgLeuProIleVala 280
DB 844 AACTATGACAGAGCATTTGGACAGTGGACACCACTTCGTTGCCCAAGAAAGTG 903
QY 281 PheGluAlaAlaValaIleSerIleValaAlaIleSerThrGluIlePheProAspGly 300
DB 904 TTTGAAAGCTGACGTCAATTCATCAAGGACGCTCTCCACGAGAAAGTTCCCTGATGT 963
QY 301 PheTYrLeuGlyGluGlnLeuValaCysTYrGlnAlaGlyTYrThrProTYrAsnIlePhe 320
DB 964 TTTGTGCTGAGAGAGAGCTGTGTCTGTGACAGAGGACCACTCCCTTGAAATTTTC 1023
QY 321 ProValIleSerLeuTYrLeuMetGlyGlyValaThrAsnGlnSerPheArgIleThrIle 340
DB 1024 CCAATCATCTCACTCACTCAATGAGGTGAGTTACCAACCAAGTCTTCGACATCACATC 1083
QY 341 LeuProGlnGlnTYrLeuArgProValaGluAspValaAlaThrSerGlnAspAspCysTYr 360
DB 1084 CTTCGCCAGCAATACCTCGCGGCGAGTGGAAAGATGTGCCCACTGCCAAGCAGACTTTTC 1143
QY 361 IysPheAlaIleSerGlnSerSerThrGlyTYrValMetGlyValaIleMetGlyIle 380
DB 1144 AAGTTGCACTTCACAGTATCAGGAGCATGTTATGAGACTGTTATCATGAGGGC 1203
QY 381 PheTYrValaIlePheAspArgAlaArgIleGlyArgIleGlyPheAlaValaSerAlaCysHis 400
DB 1204 TTCTACGTTGTCTTTGATCGGGCCGAAAACGAATTGGCTTGTCTGACGCGCTTGCAT 1263
QY 401 ValHisAspGluPheArgThrAlaAlaValaGlyIleProPheValaThrLeuAspMetGlu 420
DB 1264 GTTCACGATGAGTTCAAGCAGCGGAGGTGGAAAGGCCCTTTGTCACTTGACATGAA 1323
QY 421 AspCYseGlyTYrAsnIleProGlnTYrAspGlnSerThr 433
DB 1324 GACTGTGGCTCAACAATTCACAGACAGATGAGTCAACC 1362
RESULT 9
US-10-322-684-1
; Sequence 1, Application US/10322684
; Publication No. US20030125257A1
; GENERAL INFORMATION:
; APPLICANT: Hoffmann-La Roche Inc.
; TITLE OF INVENTION: Assay and screening method for identification of inhibitors of de
; TITLE OF INVENTION: secretases
; FILE REFERENCE: Case 21066
; CURRENT APPLICATION NUMBER: US/10/322,684
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: EP01130282.5
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1542
```

```
TYPE: DNA
ORGANISM: Homo sapiens
US-10-322-684-1
Alignment Scores:
Pred. No.: 8,886-269 Length: 1542
Score: 2267.00 Matches: 429
Percent Similarity: 99.31% Conservative: 1
Best Local Similarity: 99.08% Mismatches: 3
Query Match: 99.08% Indels: 0
DB: 15 Gaps: 0
US-10-726-967a-84 (1-433) x US-10-322-684-1 (1-1542)
QY 1 ThrGlnHisGlyTYrIleArgLeuProLeuArgSerGlyIleLeuGlyValaProLeuGlyLeu 20
DB 79 ACCCGACACGGGATCCGACTGCGACAGCGGAGCGGAGGTGCACCTTGGAACATG 138
QY 21 GluIleAsnLeuGluThrAspGluGluProGluGluProGlyArgArgGlySerPheVal 40
DB 139 CGGCTGCCCGGAGACCGACGAAAGCCGAGAGCCCGCGGAGGCGAGCTTGTG 198
QY 41 GluMetValaAspAsnLeuArgGlyIleSerGlyGlnGlyTYrTYrValaGluMetThrVal 60
DB 199 GAGATGTGGAACAACCTGAGGGGCAAGTCGGGGCAGGCTACTACGTGAGATGACCGTG 258
QY 61 GlySerProProGlnThrLeuAsnIleLeuValaAspThrGlySerSerAsnPheAlaVal 80
DB 259 GGCACGCCCCCGACAGCGCTCAACATCCGTGGATACAGGAGCAAGTAACTTGGCAGTG 318
QY 81 GlyAlaAlaProHisProPheLeuHisArgTYrTYrGlnArgGlnLeuSerSerThrTYr 100
DB 319 GGATGTGCCCCCACCCTTCTCTGATGCTACTACAGAGGACAGTCCAGCACATAC 378
QY 101 ArgAspLeuArgIleGlyValaTYrValaProTYrThrGlnGlyIleTYrGluGluLeu 120
DB 379 CGGAGCTCCGGAAGGTGTATATGATGCTTACCCAGGAGCAAGTGGAAAGGGAGCTG 438
QY 121 GlyTYrAspLeuValaSerIleProHisGlyProAsnValaThrValArgAlaAsnIleAla 140
DB 439 GGCACCGACCTGTAAAGCATCCCGCATGCGCCCAAGTCACTGTGTGCCAATGTCT 498
QY 141 AlaIleThrGlnSerAspIlePhePheIleAsnGlySerAsnTYrGluIleLeuGly 160
DB 499 GCCATCACTGAATACAGACAAGTTCTTCAACAGGCTCCAACTGGAGAGCATTCGGGG 558
QY 161 LeuAlaTYrAlaGluIleAlaArgProAspAspSerLeuGluProPhePheSerLeu 180
DB 559 CTGGCTTAAGCTGAGATTCGACGGCTGACGACTCCCTGAGCCTTTCTTGACTCTCG 618
QY 181 ValIleValThrHisValProAsnLeuPheSerLeuGluLeuCyseGlyValaGlyPhePro 200
DB 619 GTAAGACAGACCACCTTCCCAACCTCTCTCCCTGACACTTGTGTGCTGGCTTCCCC 678
QY 201 LeuAsnGlnSerGluValLeuAlaSerValaGlyGlySerMetIleIleGlyGlyIleAsp 220
DB 679 CTCACACGACTGAGAGTGTGGCTCTGTGAGGAGGATATATCATGAGAGTATGAG 738
QY 221 HisSerLeuTYrThrGlySerLeuTYrTYrThrProIleArgArgGluTYrTYrGlu 240
DB 739 CACTCGCTTAACAGCAGGAGCTCTGTGTATACCCATCCGGGAGGTGATTTATGAG 798
QY 241 ValIleIleValArgValaGluIleAsnGlyGlnAspLeuLeuMetAspCysIleGlyIle 260
DB 799 GTGATCATTTGTGGGGTGGAGATCATGACAGATCTGAAAATGAGACTGCAAGAGTAC 858
QY 261 AsnTYrAspIleSerIleValaAspSerGlyTYrThrAsnLeuArgLeuProIleVala 280
DB 859 AACTATGACAGAGCATTTGGACAGTGGACACCACTTCGTTGCCCAAGAAAGTG 918
QY 281 PheGluAlaAlaValaIleSerIleValaAlaIleSerThrGluIlePheProAspGly 300
DB 919 TTTGAAAGCTGACGTCAATTCATCAAGGAGGCTCTCCACGAGAAAGTTCCCTGATGT 978
```

QY 301 PheTrpLeuGlyGluGlnLeuValCysTrpGlnAlaGlyThrThrProTrpAsnIlePhe 320  
Db 979 TTCTGGCTAGAGAGAGAGCTGGTGTGTGGCAAGACGACACCCCTTGGAAACATTTTC 1038  
QY 321 ProValIleSerLeuTyrLeuMetGlyGluValThrAsnGlnSerPheArgIleThrIle 340  
Db 1039 CCAATCATCTCACTCTACTTAATGGGTAGGTTACCAACCAAGTCTTCCGATCAACCATC 1098  
QY 341 LeuProGlnGlnTyrLeuArgProValGluAspValAlaThrSerGlnAspAspCysTyr 360  
Db 1099 CTTCCGACGACATACCTCGGCGACGTGGAGATGTGGCCACGTCACCAAGACGATGTAC 1158  
QY 361 LysPheAlaIleSerGlnSerSerThrGlyThrValMetGlyAlaValIleMetGluGly 380  
Db 1159 AAGTTGGCATCTCACTCACTATCCAGGCACTGTATAGGAGCTGTATCATGTGAAGGC 1218  
QY 381 PheTyrValValPheAspArgAlaArgLysArgIleGlyPheAlaValSerAlaCysHis 400  
Db 1219 TTCTACGTTGTCTTGTGATCGGGCCGAAACGAATTTGGCTTTGCTGTCAAGCGCTTGCAT 1278  
QY 401 ValHisAspGluPheArgThrPheValAlaGluGlyProPheValThrLeuAspMetGlu 420  
Db 1279 GTGCACATGAGTTCAAGACGCGACGCGTGGAGGCCCTTTTGTCACTTGGACATGGA 1338  
QY 421 AspCysGlyTyrAsnIleProGlnThrAspGluSerThr 433  
Db 1339 GACTGTGGCTACACATTCGACAGACGATGAGTCAACC 1377

RESULT 10  
US-10-652-927-3  
Sequence 3, Application US/10652927  
Publication No. US20040043408A1  
GENERAL INFORMATION:  
APPLICANT: Gurney et al.  
TITLE OF INVENTION: Alzheimer's Disease Secretase, APP Substrates Therefor and Uses  
FILE REFERENCE: 2915/6280N3  
CURRENT APPLICATION NUMBER: US/10/652,927  
CURRENT FILING DATE: 2003-08-29  
PRIOR APPLICATION NUMBER: 09/794,925  
PRIOR FILING DATE: 2001-02-27  
PRIOR APPLICATION NUMBER: 09/416,901  
PRIOR FILING DATE: 1999-10-13  
PRIOR APPLICATION NUMBER: 60/155,493  
PRIOR FILING DATE: 1999-09-23  
PRIOR APPLICATION NUMBER: 09/404,133  
PRIOR FILING DATE: 1999-09-23  
PRIOR APPLICATION NUMBER: PCT/US99/20881  
PRIOR FILING DATE: 1999-09-23  
PRIOR APPLICATION NUMBER: 60/101,594  
PRIOR FILING DATE: 1998-09-24  
NUMBER OF SEQ ID NOS: 74  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 3  
LENGTH: 2070  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-652-927-3

Alignment Scores:  
Pred. No.: 1,39e-268 Length: 2070  
Score: 2267.00 Matches: 429  
Percent Similarity: 99.31% Conservative: 1  
Best Local Similarity: 99.08% Mismatches: 3  
Query Match: 99.08% Indels: 0  
DB: 18 Gaps: 0

US-10-726-967a-84 (1-433) x US-10-652-927-3 (1-2070)

QY 1 ThrGlnHisGlyIleArgLeuProLeuArgSerGlyLeuGlyValAlaProLeuGlyLeu 20  
Db 64 ACCGACGACGGGATCCGGCTCCCTCGCGACGGGCGCTGGGGGGCGCCCCCTTGGGCGCTG 123

QY 21 GluIleAsnLeuGluThrAspGluGluProGluGluProGlyValArgAlaGlySerPheVal 40  
Db 124 CGGCTCCCCGGGAGACGACGACGAGAGGCCGAGAGGCCGCGGAGGGGAGCGCTTGTG 183  
QY 41 GluMetValAspAsnLeuArgGlyLysSerGlyGlnGlyTyrTyrValGluMetThrVal 60  
Db 184 GAGATGTGTGACAACTGAGGGGCAAGTGGGGGACGGGCTACTACGTGAGATACCGCTG 243  
QY 61 GlySerProProGlnThrLeuAsnIleLeuValAspThrGlySerSerAsnPheAlaVal 80  
Db 244 GCGAGCCCCGCGACACCTCAACATCTGTGTGATACAGGACACAGTAACCTTGTGACGTG 303  
QY 81 GlyAlaAlaProHisProPheLeuHisArgTyrTyrGlnArgGlnLeuSerSerThrTyr 100  
Db 304 GGTGTGCCCCCGACCCCTTCTGATCGTACTACAGAGGACGCTGTCCAGCACATAC 363  
QY 101 ArgAspLeuArgLysGlyValTyrValProTyrThrGlnGlyLysTrpGluGlyLeu 120  
Db 364 CGGACCTCCGGAAGGGGTGTATGTGCCCTACACCGAGGCAAGTGGGAAGGGAGCTG 423  
QY 121 GlyThrAspLeuValSerIleProHisGlyProAsnValThrValArgAlaAsnIleAla 140  
Db 424 GGCACCGACTGTGTAGCATCCCAATGCCCAAGCTCATCTGTGCGTCCACATTTGCT 483  
QY 141 AlaIleThrGluSerAspLysPhePheIleAsnGlySerAsnTrpGluGlyIleLeuGly 160  
Db 484 GCCATCACTGAAATGACAAAGTTCTTCATCAACGGCTCCAACTGGGAAGGCATCTCGGG 543  
QY 161 LeuAlaTyrAlaGluIleAlaArgProAspAspSerLeuGluProPheAspSerLeu 180  
Db 544 CTGGCTATGTGATGATGGCAGGCTGACGACATCCCTGTGAGGCTTTTGTGCTCGTCTG 603  
QY 181 ValLysGlnThrHisValProAsnLeuPheSerLeuGlnLeuCysGlyAlaGlyPhePro 200  
Db 604 GTAAAGCAGACCCACGTTCCCAACTCTTCTCCCTGACACTTGTGTGCTGCTGCTTCCC 663  
QY 201 LeuAsnGlnSerGluValLeuAlaSerValGlyLysMetIleIleGlyGlyLeuAsp 220  
Db 664 CTCAACGAGTCTAAAGTGTGGCTGTGTGCGGAGGACATGATCATTTGAGATTCGAC 723  
QY 221 HisSerLeuTyrThrGlySerLeuTyrTyrThrProIleArgArgGluTyrTyrGlu 240  
Db 724 CACTCGGTGTACACAGGAGTCTCTGTATACCCATCCGCGGAGTGTATATGAG 783  
QY 241 ValIleIleValAlaArgValGluIleAsnGlyGlnAspLeuLysMetAspCysLysGluTyr 260  
Db 784 GTCATATTGTGGGTGTGAGATCAATGACAGGATCTGAAATGAGACTGCAGAGAGTAC 843  
QY 261 AsnTyrAspLysSerIleValAspSerGlyThrThrAsnLeuArgLeuProLysVal 280  
Db 844 AACTATGACAAAGCATTTGTGACAGTGGACCAACCACTTCGTTTCCCAAGAAATGTG 903  
QY 281 PheGluAlaAlaValLysSerIleLysAlaAlaSerSerThrGluLysPheProAspGly 300  
Db 904 TTGGAAGTGTGCAATCAATCCATCAAGGAGCTCTCTCAAGGAAATTCCTGATGCT 963  
QY 301 PheTrpLeuGlyGluGlnLeuValCysTrpGlnAlaGlyThrThrProTrpAsnIlePhe 320  
Db 964 TTCTGGCTAGAGAGAGCTGGTGTGTGGCAAGACGACACACCCCTTGGAAACATTTTC 1023  
QY 321 ProValIleSerLeuTyrLeuMetGlyGluValThrAsnGlnSerPheArgIleThrIle 340  
Db 1024 CCAATCATCTCACTCTACTTAATGGGTAGGTTACCAACCAAGTCTTCCGATCAACCATC 1083  
QY 341 LeuProGlnGlnTyrLeuArgProValGluAspValAlaThrSerGlnAspAspCysTyr 360  
Db 1084 CTTCCGACGACATACCTCGGCGACGTGGAGATGTGGCCACGTCACCAAGACGATGTAC 1143  
QY 361 LysPheAlaIleSerGlnSerSerThrGlyThrValMetGlyAlaValIleMetGluGly 380  
Db 1144 AAGTTGGCATCTCACTCACTATCCAGGCACTGTATAGGAGCTGTATCATGAGAGGC 1203









```

; LENGTH: 2070
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-940-867-3

Alignment Scores:
Pred. No.: 1,39e-268 Length: 2070
Score: 2267.00 Matches: 429
Percent Similarity: 99.31% Conservative: 1
Best Local Similarity: 99.08% Mismatches: 3
Query Match: 99.08% Indels: 0
D: 21 Gaps: 0

US-10-726-967a-84 (1-433) x US-10-940-867-3 (1-2070)

QY 1 ThrGlnHISGLYILEArgLeuProLeuAArgSerGlyLeuGlyGlyValAProLeuGlyLeu 20
Db ACCGACGACGAGGATCCGCGCTGCGCCGACGCGCTGGGGGGCCCCCTGGGGCTG 123
QY 21 GlnLeuAsnLeuGlyThrAspGluGluProGluGluProGlyArgArgGlySerPheVal 40
Db CGGCTGCGCGGAGAGCCGACGAGAGCCGAGAGACCCGCGCGGAGGGGCAAGCTTGTG 183
QY 41 GlnMetValAspAsnLeuArgGlyLysSerGlyGlnGlyTyrTyrValGlnMetThrVal 60
Db GAGATGCTGACAACTGAGGGGCAAGTGGGGCAAGGCTACTACCTGAGATGACCGTG 243
QY 61 GlySerProGluThrLeuAsnIleLeuValAspThrGlySerSerAsnPheAlaVal 80
Db GGACGACCCCGACGAGCGCTCAACATCTGGTGGATACAGGACGACGATCTTGGACGTG 303
QY 81 GlyAlaAlaProHisAspPheLeuHisArgTyrTyrGlnArgGlnLeuSerSerThrTyr 100
Db GGGTGGCGCCCGACCCCTTCCGACATCGCTACTACAGAGGACGCTGCCAGCATATAC 363
QY 101 ArgAspLeuArgLysGlyValTyrValProTyrThrGlnGlyLysArgGlyGlnLeu 120
Db CGGACCTCCGAGAGGGGTGTGTATGTGCCCTTACACCGAGGGCAAGTGGAGGGGAGCTG 423
QY 121 GlyThrAspLeuValSerIleProHisGlyProAsnValThrValArgAlaAsnIleAla 140
Db GGCACCGACCTGGTAAAGCATCCCCCAATGGCCCAAGTCATCTGGCTGCCAATTCGCT 483
QY 141 AlaIleThrGlySerAspLysPhePheIleAsnGlySerAsnTrpGluGlyIleLeuGly 160
Db GCCATCACTGAATCAGACAAAGTCTTTCATCAAGGCTCCCAACTGGGAAGCATCTGGG 543
QY 161 LeuAlaTyrAlaGluIleAlaArgProAspAspSerLeuGluProPhePheAspSerLeu 180
Db CTGGCTTATGTGATGATTCGACAGGCTGACGACTCCCTGGAGCCCTTCTTGTGACTCTG 603
QY 181 ValIleGlnThrHisValProAsnLeuPheSerLeuGlnLeuGlyGlyAlaGlyPhePro 200
Db GTAAACAGACCAACGCTCCCAACCTTCTCCCTGACGTTGTGGTCTGGCTTCCCC 663
QY 201 LeuAsnGlnSerGlyLeuAlaLeuAlaSerValGlyGlySerMetIleIleGlyGlyIleAsp 220
Db CTCAACCAAGCTGAAGTCTGGCTCTGTGCGAGGAGCATGATCATGAGGATTCGAC 723
QY 221 HisSerLeuTyrThrGlySerLeuTyrTyrThrProIleArgArgGluTyrTyrGln 240
Db CACTCGCTGACACAGGCACTCTCTGTATACACCCATCCGGCGGGGATGATTTATGAG 783
QY 241 ValIleIleValArgValGluIleAsnGlyGlnAspLeuLysMetAspCysLysGluTyr 260
Db GTATCTCTTGTGGGGGAGATCAATGACAGGATCTGAATATGACTGCAAGAGATAC 843
QY 261 AsnTyrAspLysSerIleValAspSerGlyThrThrAsnLeuArgLeuProLysVal 280
Db AACTATGACAAAGCATTTGTGACAGTGGACACCAACCTTGCTTGGCCCAAGAAAGTG 903
QY 281 PheGluAlaAlaValLysSerIleLysAlaAlaSerSerThrGlyLysPheProAspGly 300

```

```

Db 904 TTTGAAAGTCGACGTAATTCATCAAGGACGCTCTCCACGAGAAAGTTCCCTGATGCT 963
QY 301 PheTrpLeuGlyGlnGlnLeuValCysTrpGlnAlaGlyThrThrProTrpAsnIlePhe 320
Db TTCTGGCTAGGAGAGACACTGGTGTGCTGGCAAGGACGACACACCTTGGAAACATTTTC 1023
QY 321 ProValIleSerLeuTyrLeuMetGlyGlyValValThrAsnGlnSerPheArgIleThrIle 340
Db CGATCATCTCACTTACTTAATGGTATGAGTTACCAACCAAGTCTCTCCGATCAACATTC 1083
QY 341 LeuProGlnGlnTyrLeuArgProValGlnLysValAlaThrSerGlnAspAspCysTyr 360
Db CTTCGCGACAAATACCTCGGCGCAAGTGGCCAGCTCCCAAGACGACTGTATAC 1143
QY 361 LysPheAlaIleSerGlnSerSerThrGlyThrValMetGlyAlaValIleMetGlnGly 380
Db AAGTTGGCATCTCAAGATCATCCACGGGCACTGTATAGGAGCTGTATATCATGGAGGCG 1203
QY 381 PheTyrValValPheAspArgAlaArgLysArgIleGlyPheAlaValSerAlaCysHis 400
Db TTCTACGTTGTCTTGAATCGGCGCCGAAACGAAATGGCTTTGCTGTCAGCGCTTGCCAT 1263
QY 401 ValHisAspGluPheArgThrAlaAlaValGlnGlyProPheValThrLeuAspMetGlu 420
Db GTGCAAGATGATTCAGAGCGGCAAGGTCAGGAGGCTTTGTGTCACCTTGGACATGMA 1323
QY 421 AspCysGlyTyrAsnIleProGlnThrAspGluSerThr 433
Db 1324 GACTGTGGCTACAACTTCCACAGACAGATGATGATCAACC 1362

RESULT 14
US-11-089-918-44
; Sequence 44, Application US/11089918
; Publication No. US20050164327A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, John P.
; APPLICANT: Basl, Gurigdal
; APPLICANT: Doane, Minh Tam
; APPLICANT: Frigon, Normand
; APPLICANT: John, Varghese
; APPLICANT: Power, Michael
; APPLICANT: Sinha, Sukanto
; APPLICANT: Tatsuno, Gwen
; APPLICANT: Tung, Jay
; APPLICANT: Wang, Shuwen
; APPLICANT: McConlogue, Lisa
; TITLE OF INVENTION: Beta-Secretase Enzyme Compositions and
; TITLE OF INVENTION: Methods
; FILE REFERENCE: 228-US-NEW2C9
; CURRENT APPLICATION NUMBER: US/11/089,918
; PRIOR FILING DATE: 2005-03-25
; PRIOR APPLICATION NUMBER: US 09/723,722
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 09/501,708
; PRIOR FILING DATE: 2000-02-10
; PRIOR APPLICATION NUMBER: US 09/471,669
; PRIOR FILING DATE: 1999-12-24
; PRIOR APPLICATION NUMBER: 60/119,571
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: 60/139,172
; PRIOR FILING DATE: 1999-06-15
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 2348
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-089-918-44

Alignment Scores:
Pred. No.: 1.69e-268 Length: 2348
Score: 2267.00 Matches: 429
Percent Similarity: 99.31% Conservative: 1

```

Best Local Similarity:	99.08%	Mismatches:	3
Query Match:	99.08%	Indels:	0
DB:	24	Gaps:	0
US-10-726-967A-84 (1-433) x US-11-089-918-44 (1-2348)			
QY	1	ThrGlnHisGlyIleArgLeuProLeuArgSerGlyLeuGlyValAlaProLeuGlyLeu	20
DB	369	ACCCAGACAGCGCATCCGGCTGCCCTCCGACAGCGGCTGGGGGGCCCCCCCCCGGGGCTG	428
QY	21	GluIleHisLeuGlyIleThrArgGlyIleProGlyIleProGlyIleValArgGlySerPheVal	40
DB	429	CGCGTCCCGCGGAGCCGACGAAAGCCCGAGACCCCGCGGCGGGCGGCTTGTG	488
QY	41	GluMetValAspAsnLeuArgGlyIleSerGlyIleGlyIleValGluMetThrVal	60
DB	489	GAATGTGGACAACTCGAGGGGCAAGTCGGGGCAAGGCTACTACATCGAGATGACCGTG	548
QY	61	GlySerProProGlnThrLeuAsnIleLeuValAspThrGlySerSerAsnPheAlaVal	80
DB	549	GGCAGCCCCCGACAGCGCTCAACATCTGTGTGATACAGGCGACGATACCTTGCAGTG	608
QY	81	GlyAlaAlaProHisProPheLeuHisArgIleValArgGlyIleSerSerThrIle	100
DB	609	GGTGTGCCCCCGACCCCTTCTCGCATCGCTACTACAGAGGCGAGCTTCCAGCACATAC	668
QY	101	ArgAspLeuArgGlyValIleValIleProIleThrGlnGlyIleValIleGlyIleVal	120
DB	669	CGGACCTCCGAAAGGTGTGTATGTGCTTACCCCTACACCAAGGCAAGTGGAGGGAGCTG	728
QY	121	GlyThrAspLeuValSerIleProHisGlyIleProAsnValThrValArgAlaAsnIleAla	140
DB	729	GGCAGCGACCTGGTATGATCCCGACATGGCCCCCAAGCATGTGTGCTGCCAATGTGT	788
QY	141	AlaIleThrGlnSerAspIlePhePheIleHisGlySerAsnIleGlyIleGlyIle	160
DB	789	GGCATCTACGAAATCAACAAGTCTTCTCATCAACGCTCCAACTGGAGAGGCACTCTGGGG	848
QY	161	LeuAlaIleValAlaGluIleAlaArgProAspAspSerLeuGlyProPhePheAspSerLeu	180
DB	849	CTGGCTTATGCTGAAATTCGACGGCTTGACGACTCCCTGGAGCTTCTTTTACCTCTG	908
QY	181	ValIleGlnThrHisIleValProAsnLeuPheSerLeuGlnLeuGlyValAlaGlyPhePro	200
DB	909	GTAAGACAGACCCAGCTTCCCAACCTCTTCCCTGACAGCTTGTGTGCTGCTGCC	968
QY	201	LeuAsnGlnSerGlyValLeuAlaSerValGlyIleSerMetIleIleGlyIleValAsp	220
DB	969	CTCAACCAAGTGAAGTGTGCTCTTGTGCGAGGAGGATGATATGAGAGTATCGAC	1028
QY	221	HisSerLeuIleThrGlySerLeuIleValIleValIleValIleValIleValIleVal	240
DB	1029	CACCTGCTGTACACAGCGAGTCTCTGTGTATACCCATCCGGCGGAGGTGATTTATGAG	1088
QY	241	ValIleIleValAlaArgValGluIleHisGlyIleAspLeuValMetAspCysIleValIle	260
DB	1089	GTGATCATTTGTGCGGTGGAGATCAATGACAGATCTGAAATAAGACTGCAAGAGATAC	1148
QY	261	AsnIleValAspIleSerIleValAspSerGlyIleThrAsnLeuArgLeuProIleValIle	280
DB	1149	AACTATGACAAAGCATTTGTGACAGTGCACACCAACCTTCTGTTCCCAAGAAAGTG	1208
QY	281	PheGluAlaAlaValIleSerIleValAlaIleValIleValIleValIleValIleVal	300
DB	1209	TTTGAAGCTGCGATCAATCCATCAAGCGAGCTCTTCCACGGAAGAAGTCTCTATGCT	1268
QY	301	PheIlePheGlyIleGlnIleValIleValIleValIleValIleValIleValIleVal	320
DB	1269	TTCTGGCTAGAGACAGCTGTGTGTGCTGGCAAGGACACCAACCTTGGAAACTTTTC	1328
QY	321	ProValIleSerLeuIleValIleValIleValIleValIleValIleValIleValIle	340
DB	1329	CGATCATCTACCTCACTAATGAGGTGATTCACCAACAGTCTTCCGATCACCATC	1388

QY	341	LeuProGlnGlnIleValIleValIleValIleValIleValIleValIleValIleVal	360
DB	1389	CTTCCGCGCAATATCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1448
QY	361	LysPheAlaIleSerGlnSerSerThrGlyIleValIleValIleValIleValIleVal	380
DB	1449	AAATTTGCACTACATCACTACATCACTACATCACTACATCACTACATCACTACAT	1508
QY	381	PheIleValIleValIleValIleValIleValIleValIleValIleValIleValIle	400
DB	1509	TTCTACGTTGTCTTGTGATCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1568
QY	401	ValIleAspGlyIlePheArgThrAlaAlaValGluIleProPheValIleLeuAspMetGly	420
DB	1569	GTGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1628
QY	421	AspCysGlyIleValIleValIleValIleValIleValIleValIleValIleValIle	433
DB	1629	GACTGTGCTTCAACATTTCCACAGACAGATGATGATGATGATGATGATGATGATGAT	1667
RESULT 15			
US-11-090-866-44			
Sequence 44, Application US/11090866			
Publication No. US20050164294A1			
GENERAL INFORMATION:			
APPLICANT: Anderson, John P.			
APPLICANT: Basal, Gurigdal			
APPLICANT: Doane, Minh Tam			
APPLICANT: Frigon, Normand			
APPLICANT: John, Varghese			
APPLICANT: Power, Michael			
APPLICANT: Sinha, Sukanto			
APPLICANT: Tatsuno, Gwen			
APPLICANT: Tung, Jay			
APPLICANT: Wang, Shuwen			
APPLICANT: McConlogue, Lisa			
TITLE OF INVENTION: Beta-Secretase Enzyme Compositions and			
TITLE OF INVENTION: Methode			
FILE REFERENCE: 228-US-NEW2C12			
CURRENT FILING DATE: 2005-03-25			
PRIOR APPLICATION NUMBER: US/11/090,866			
PRIOR FILING DATE: 2000-11-28			
PRIOR APPLICATION NUMBER: US 09/501,708			
PRIOR FILING DATE: 2000-02-10			
PRIOR APPLICATION NUMBER: US 09/471,669			
PRIOR FILING DATE: 1999-12-24			
PRIOR APPLICATION NUMBER: 60/119,571			
PRIOR FILING DATE: 1999-02-10			
PRIOR APPLICATION NUMBER: 60/139,172			
PRIOR FILING DATE: 1999-06-15			
NUMBER OF SEQ ID NOS: 104			
SOFTWARE: FastSeq for Windows Version 4.0			
SEQ ID NO 44			
LENGTH: 2348			
TYPE: DNA			
ORGANISM: Homo sapiens			
US-11-090-866-44			
Alignment Scores:			
Pred. No.:	1,69e-268	Length:	2348
Score:	2267.00	Matches:	429
Percent Similarity:	99.31%	Conservative:	1
Best Local Similarity:	99.08%	Mismatches:	3
Query Match:	99.08%	Indels:	0
DB:	24	Gaps:	0
US-10-726-967A-84 (1-433) x US-11-090-866-44 (1-2348)			
QY	1	ThrGlnHisGlyIleArgLeuProLeuArgSerGlyLeuGlyValAlaProLeuGlyLeu	20
DB	369	ACCCAGACAGCGCATCCGGCTGCCCTCCGACAGCGGCTGGGGGGCCCCCCCCCGGGGCTG	428

Qy	21	GIuileAhneJuglUthrrAspGluJuglUProglUglUProglUthAryAGyAgglYserPheVal	40
Db	429	CGGCTGCCCCGGGAGACCCAGCAAGAGCCCGAGAGCCCGGCGGAGGGGACGCTTTGTG	488
Qy	41	GIuueValAAspAenLeuArgIlyVserGIyGlnGIyTyTyValGIuueThrVal	60
Db	489	GAGATGGTGACAACTGAGGGGCAAGTCGGGGGACAGGGCTACTACGTGAGATGACCGTG	548
Qy	61	GIyserProProGIInThrrLeuAanIIeLeuValAspThrGIySerSerAanPheAlaVal	80
Db	549	GCGAGCCCCCGCAGACGCTCAACACTCTGTGGATACAGGACAGCAATTACTTTGCACTG	608
Qy	81	GIyAlaAlaIProHIsProPheLeuHIsArgTyTyGIyGlnArgIInLeuSerSerThrTy	100
Db	609	GGTGTGCCCCCACCCTTCCTGCAATGCTACTACCAAGGACGAGCTGTCCAGACATAC	668
Qy	101	ArgAspLeuArglyGIyValTyTyValProTyThrGlnGlylySTrPGIuGIyGluLeu	120
Db	669	CGGACCTCCGGAAGGGGTGTATGTGCGCTACACCAAGGCGCAAGTGGGAAGGGACCTG	728
Qy	121	GIyThrAspLeuValSerIleProHIsGIyProAanValThrValArgAlaAanIIeAla	140
Db	729	GGCACCGACCTGGTAAAGCATCCCAATGGCCCCCAAGCTCACTGTGCTGCCAATTTGCT	788
Qy	141	AlaIleThrGIuSerAspLyPhePheIIeAanGlySerAanTrpGIuGlyIleLeuGly	160
Db	789	GCATCACTGAATCAGACAAAGTTCTTCTATCAACGGCTTCAACTGGGAAGGCATCTGGGG	848
Qy	161	LeuAlaTyAlaGIuIIeAlaArgProAspAspSerLeuGIuProPhePheAspSerLeu	180
Db	849	CTGGCGTATGTGTGAGATTGCCAGGCTGTACACACTCCCTGGAGCCCTTTCTTGACTCTGTG	908
Qy	181	ValIyGlnThrrHIsValProAsnLeuPheSerLeuGIInLeuCyGIyAlaGIyPhePro	200
Db	909	GTTAAAGCAGACCCAGCTTCCCAACTCTTCTCCCTGACAGCTTTGTGGTGGCTTCCCC	968
Qy	201	LeuAanGlnSerGIuValIleuAlaSerValGIyGlySerMetIleIleGlyGlyIleAsp	220
Db	969	CTCAACCGCTGAAAGTGTCTGGCTCTGTCCGAGGGAGACAGATCAATTGGAGGATGTGAC	1028
Qy	221	HIsSerLeuTyTrpThrGIySerLeuTrpTyTrpThrProIleArgAGGIIuTrpTyTyGIu	240
Db	1029	CACCTGCTGTAACAGGACAGTCTGTGGTATACACCACTCCGCGGAGTGGTATTATGAG	1088
Qy	241	ValIleIleValArgValGIuIIeAanGlyGlnAspLeuYsMetAspCyLyGIyGlyTr	260
Db	1089	GTCATCATTTGGCGGGTGGAGATCAATGACAGAGATCTGAATATGCACTGCACAGAGATAC	1148
Qy	261	AanTyAspLySerIleValAspSerGIyThrThrAanLeuAglLeuProLyLyVal	280
Db	1149	AACATATGACAAAGACATTTGTGGACAGTGGACACACCAACTTGTCTTGGCCCAAGAAATG	1208
Qy	281	PheGIuAlaAlaValIySerIleValAlaIAserSerThrGIuLyPheProAspGIy	300
Db	1209	TTTGAAGCTGCAGTCAATTCATCAAGGCAAGCTCTCTCAAGGAAGATTCCTCGATGGT	1268
Qy	301	PheTrpLeuGIyGluGlnLeuValCySTrPGIuAlaGIyThrTrpProTrpAanIIePhe	320
Db	1269	TTCTGGCTAAGAGACAGCTGTGTGTGGCAAGGACACACACCCCTTGGAAACATTTTC	1328
Qy	321	ProValIleSerLeuTyTrpLeuMetGIyGluValThrAanGlnSerPheArgIIeThrIle	340
Db	1329	CCAGTCATCTCACTTACTTAATGGGTGAAGTTACCAACAGTCTTCCGATCACCATC	1388
Qy	341	IeuProGIInTyTrpLeuArgProValGIuAspValAlaTrpSerGIInAspAspCySTr	360
Db	1389	CTTCCGACAGCATTACTCGGGCCAGTGGAAAGATGTGGCCAGTCCCAAGACAGACTGTAC	1448
Qy	361	LySPheAlaIleSerGIInSerSerThrGIyTrpValMetGIyAlaValIIeMetGIuGly	380
Db	1449	AAGTTTGCATCTCAACATCATCAACGGGCACTGTTATGGAGCTGTTATCATGGAGGAC	1508

Qy	381	PheYtrValValAlaPheAspArgAlaArgLysArgIleGlyIlePheAlaValSerAlaCysHis	400
Db	1509	TTCTACAGTTGCTCTTGTGATCGGAGCCGAAAGCAATTGGCTTGTGTCAGCCCTTGGCAT	1566
Qy	401	ValHisAspGluPheArgThrAlaAlaValGluGlyProPheValThrLeuAspMetGlu	420
Db	1569	GTGCACGATGAGTTCAAGCAGCGCAGCGGTGGAAAGCCCTTTGTGCACCTTGACATGGA	1628
Qy	421	AspCysGluYTYraenIleProGlnThrAspGluSerThr	433
Db	1629	GACTGTGGCTCAACATTCCACAGACAGATGATCAACCC	1667

Search completed: August 22, 2005, 05:30:02  
Job time : 1246 secs

Search completed: August 22, 2005, 05:30:02  
Job time : 1246 secs

**This Page Blank (uspto)**